现代数学译丛 2

金融数学引论

——从风险管理到期权定价

〔美〕Steven Roman 著

邓欣雨 译



(O-2964.0101)


ISBN 978-7-03-020744-9

销售分类建议：高等数学；金融

定 价：63.00 元

现代数学译丛 2

金融数学引论

——从风险管理到期权定价

〔美〕 Steven Roman 著

邓欣雨 译

科学出版社

北京

图字: 01-2007-3531 号

内 容 简 介

本书介绍投资组合风险管理和期权定价等金融数学的基本知识，主要包括资本资产定价模型(CAPM)、Black-Scholes 期权定价模型以及未定权益定价中常用的无套利原理和鞅方法. 每章结合实例解释基本概念，并配有一定量的习题.

本书适合作为高等院校数学、金融或经济学专业的高年级本科生或研究生教材，也可作为金融证券类从业人员的参考书.

图书在版编目(CIP)数据

金融数学引论:从风险管理到期权定价/(美) Steven Roman 著;
邓欣雨译. —北京: 科学出版社, 2008
(现代数学译丛; 2)
ISBN 978-7-03-020744-9

I. 金… II. ①诺… ②邓… III. 金融-经济数学 IV. F830

中国版本图书馆 CIP 数据核字(2007) 第 196408 号

责任编辑: 赵彦超 / 责任校对: 陈玉凤
责任印制: 赵德静 / 封面设计: 王 浩

科 学 出 版 社 出版
北京东黄城根北街 16 号
邮政编码: 100717
http://www.sciencep.com

源海印刷有限责任公司 印刷
科学出版社发行 各地新华书店经销
*
2008 年 1 月第 一 版 开本: B5(720×1000)
2008 年 1 月第一次印刷 印张: 18 1/2
印数: 1—3 000 字数: 324 000

定价: 63.00 元

(如有印装质量问题, 我社负责调换〈明辉〉)

献给多纳 (Donna)

前言

本书覆盖了金融数学的两大主要领域. 其一是投资组合的风险管理, 在资本资产定价模型处达到最高点; 其二是资产定价理论, 在布莱克 – 舒尔斯期权定价公式处达到最高点. 我们只用一章来讨论投资组合的风险管理. 剩下的部分专注于资产定价模型的研究, 这是当今令人非常感兴趣和研究较多的主题.

本书适合数学、金融或者经济系的高年级本科生或者低年级研究生阅读. 因此, 在本书中没有用到测度论知识.

我认识到本书的读者可能具有不同的背景. 一方面, 数学系的学生在数学上的思考可能有很好的准备, 但是在金融方面 (投资组合、股票期权、远期合约等) 就没有很好的准备了. 另一方面, 金融系和经济系的学生可能精通金融方面的话题, 但是没有数学系学生那么好的数学思维.

既然本书的主题是金融数学, 所以我没有以任何方式来冲淡数学知识 (当然是本书的相应水平). 这就是说, 一方面竭力在相应水平上保持数学上的严谨, 另一方面假设读者没有金融背景, 介绍了一些必要的金融背景知识 (股票期权和远期合约).

我也努力使本书尽量在数学上独立. 除了一定的数学水平外, 掌握一年级的线性代数已经足够. 特别地, 读者应该了解矩阵代数、向量空间和线性变换的核及值域这些概念. 虽然在风险管理方面的一些证明中用到了拉格朗日乘子, 但是如果你喜欢的话, 这些证明可以省略.

当然, 概率论也出现在金融数学领域. 本书在这方面是独立的. 全书有几章专门讲述概率论知识. 该想法是为了提供一些需要了解的必要理论. 这样, 如果读者不打算学习连续情形下的定价理论, 就不需要涉及与连续概率有关的知识.

本书的安排如下. 第 1 章主要讨论离散概率的基础知识. 此章包括诸如随机变量、独立性、期望、方差和最佳线性估计等内容. 如果读者已经学过初等概率论, 那么这一章可以当作一个回顾.

第 2 章主要讨论投资组合理论和风险管理. 主要目的是为了描述著名的资本资产定价模型 (CAPM). 此章与剩余的章节相互独立, 如果喜欢的话, 你可以跳过此章的内容.

本书剩余的章节致力于资产定价模型的讨论. 第 3 章给出了股票期权的必要背景知识. 第 4 章在无套利假设下简要地阐明了资产定价的方法, 主要通过定价标准的远期合约和讨论一些与期权定价相关的话题来说明定价方法, 例如买权和卖权的平价关系, 这牵涉到标的物相同、执行价相同且到期日也相同的买权和卖权的价格.

第 5 章继续讨论离散概率, 包含条件概率和一些较高等的概率知识, 例如样本空间的划分、随机变量、(关于样本空间一个划分的) 条件期望、随机过程和鞅论. 这些都是在离散情形下进行的, 对有些学生来说可能是第一次接触到.

有了第 5 章的概率知识, 读者就已经为处理第 6 章的离散时间模型做好了准备. 第 7 章描述了考克斯 – 罗斯 – 鲁宾斯坦模型 (Cox-Ross-Rubinstein). 此章比较短, 但是介绍了漂移、波动率和随机游走等一些重要问题.

第 8 章介绍了连续概率的一些很基本的知识. 我们需要依分布收敛和中心极限定理的概念, 这样就能在每期的时间长度趋于零时, 对考克斯 – 罗斯 – 鲁宾斯坦模型取极限. 我们在第 9 章实行了极限过程, 得到了著名的布莱克 – 舒尔斯期权定价公式.

第 10 章讨论了可选停时和美式期权. 相对于前面的章节来说, 此章在数学上可能具有挑战.

最后还有两节附录, 都属于可选择的. 在附录 A 中, 我们讨论了在离散模型中对不可达到的未定权益进行定价这个问题. 这可在阅读了第 6 章后的任何时间里来阅读这些材料. 附录 B 涉及了凸性方面的背景知识, 在第 6 章中需用到它.

关于定义的注记

与数学的其他领域不同, 本书的主题 —— 金融数学, 还没有多少适合本科生阅读的著作. 更简单地说, 在金融数学方面很缺乏本科生教材.

因此, 在大学阶段很少有先例制定基本的理论, 因为在该阶段, 教育理论和直觉的运用是第一位的. 缺乏专业的术语来处理某些状况就表明了这一点.

相应地, 很少有情况使我感觉发明新的术语来代替一些特殊的概念是必要的. 可以向读者保证我没有这样做. 我发明术语不是为了任何别的原因, 只是为了便于更好地传授知识.

无论如何, 读者会碰到一些我标记了 “非标准” 的定义. 该标记是为了传达一个事实, 即这些定义可能在其他书中找不到, 也可能超出了本书的范围就不能使用.

致谢

最后, 要感谢我的学生 Lemmee Nakamura,Tristan Egualada 和 Christopher Lin, 他们耐心地听了我的预备课并对原稿提出了宝贵的意见. 当然本书的任何错误 (希望尽量少) 都是我个人的责任, 欢迎读者访问我的网站 www.romanpress.com, 可以更多地了解我的书籍或者留下评论和建议.

Steven Roman

美国加州大学欧文校区 (Irvine)

符号标记和希腊字母表

$\langle , \rangle$	$\mathbb{R}^n$ 上的内积 (点乘)
$\mathbf{1}$	单位向量 $(1, \cdots, 1)$
1_A^S	$A \subseteq S$ 的示性函数
$\mathcal{A} = \{a_1, \cdots, a_n\}$	n 种资产
C	买权价格
e_i	标准单位向量的第 i 个元素
$\mathcal{E}_P(X)$	X 关于概率测度 P 的期望
Φ_i	资产持有过程
K	执行价
μ_X	X 的期望
$\Omega = \{\omega_1, \cdots, \omega_m\}$	经济的状态
P	卖权价格
$\mathcal{P}_i = \{B_{i,1}, \cdots, B_{i,m_i}\}$	状态的划分
$\mathbb{P}$	概率测度
r	无风险利率
$\mathrm{RV}(\Omega)$	从 Ω 到 $\mathbb{R}$ 的所有随机变量形成的向量空间
$\mathrm{RV}^n(\Omega)$	从 Ω 到 $\mathbb{R}^n$ 的所有随机变量形成的向量空间
S	股票或其他资产的价格
$\sigma = (s_1, \cdots, s_m)$	状态向量
σ_X^2	X 的方差
$\sigma_{X,Y}$	X 与 Y 的协方差
Θ_i	投资组合
$\mathcal{V}_0$	初始成本函数
$\mathcal{V}_T$	损益函数

希腊字母表

A α alpha	H η eta	N ν nu	T τ tau
B β beta	Θ θ theta	Ξ ξ xi	Υ, υ upsilon
Γ γ gamma	I ι iota	O o omicron	Φ ϕ phi
Δ δ delta	K κ kappa	Π π pi	X χ chi
E ϵ epsilon	Λ λ lambda	P ρ rho	Ψ ψ psi
Z ζ zeta	M μ mu	Σ σ sigma	Ω ω omega

目 录

导　言

投资组合的风险管理

为了比其他人赚更多的钱, 风险是不可能避免的. 固然, 钱生钱这个过程可以不需承担任何风险就可以完成: 投资者所需要做的就是购买联邦国债, 一般被认为是无风险投资. 该类投资得到的收益率一般不令人满意.

现实问题是, 如果每个人都得到相同的收益率, 那么该收益仅仅是为了维持现状. 换句话说, 如果你想买一辆劳斯莱斯或者一艘游艇, 甚至是一块劳力士手表, 那么你需要比其他人赚更多的钱, 这就需要承担风险.

首要问题就是如何度量一种资产的风险. 这被证明是很简单的. 然而, 像生命的剩余部分那样, 简单的答案经常被证明是不完美的. 特别地, 不仅度量资产的风险很重要, 而且度量组合中一种资产与另一资产的交互风险也很重要. 毕竟, 最后是靠投资者的整个投资组合表现来区别各投资者的.

当然, 一般来说, 一种资产的未来回报在当前是未知的. 从概率方面说, 资产的回报是**随机变量**. 因此, 整个资产组合的回报也是随机变量. 然而, 不难发现, 通过细心组合资产可能改变组合的全部风险, 该风险甚至低于每一单个资产的风险. 这种通过资产选择降低风险的过程称为**分散化**.

从数学上讲, 被普遍接受的一种度量资产风险的好方法就是回报的方差 (或标准差). 读者可能知道, 方差 (或标准差) 衡量了随机变量可能取值的分散程度. 方差越大, 风险显著偏离平均值的这个概率值就越大. 利用相同的表示, 组合中一种资产的回报与另一资产的回报的协方差就为交互风险提供了一种很好的度量.

因此, 在投资组合管理理论中, 需要求资产回报的期望值, 也需要求它的方差以及与其他资产之间的协方差. 基于风险 - 收益分析考虑, 只有通过这些统计量才能确定将一特定资产加入到组合中去是否合适, 如我们将看到的, 一个值得注意的简单程序是由资本资产定价模型 (CAPM) 提供的.

CAPM 引出了**市场组合**概念, 市场组合是**完全分散化**的风险资产组合. 从理论上看, 该组合一定包含了所有可行的资产. 这是因为所有的投资者都想唯一地投资该组合 (与无风险资产一起), 从而任何不被市场组合包含的资产都会被忽视而消失.

从实践的角度来看, 市场组合就只是一股热空气. 另一方面, 研究表明, 可以通过有效地投资少数资产来逼近市场组合. 幸运的是, 这也部分地减轻了不良资产问题, 因为一种未能进入一投资者“市场组合”的资产可能进入另一投资者的组合.

一旦市场组合 (或者它的近似组合) 被确定, 对于理性投资者 (至少在理论上) 来说, 需要考虑的唯一问题就是如何在风险组合和无风险资产上投资. 这不是数学问题而是投资者的个人内部问题了.

期权定价模型

金融证券或**金融工具**是具有法律效力的合约, 该合约传达了所有权 (比如股票情形)、信用度 (如债券情形) 或者所有者的权力 (如股票期权情形).

一些金融证券的价值依赖另一证券的价值, 这时前一证券称为后一证券的**衍生品**, 而后一证券称为衍生品的**标的证券**. 衍生产品中最著名的例子就是普通的股票期权 (卖权和买权). 这种情况下的标的证券就是股票.

然而, 衍生产品已变得如此流行, 以致它们的标的物更多是一些奇异金融实体, 有一些都不能称为正式的金融证券, 比如利率和汇率. 可能这就是通常将标的实体简单地称为**标的物**的原因吧.

衍生品也可能基于其他衍生品而被衍生出来的. 例如, 可以交易期货合约上的期权. 因此, 某一给定的金融实体在一些情况下可能是衍生品, 在另一些情况下可能就是标的物.

确实, 投资者的商业行为就是为了赚钱, 这可通过承担风险而实现, 意思是说赌博. 与拉斯维加斯的娱乐场所总是在寻找新的赌博来增加自己的利润一样, 投资团体也总在寻找新的金融赌博机会. 这些赌博通常采用奇异衍生品的形式.

在本书中, 我们专注于简单衍生品, 主要是普通的股票期权. 我们对买卖这样的证券很感兴趣. 买方购买一种证券, 就说他在该证券上处于**多头** (long position), 卖方卖出一种证券, 就说他处于**空头** (short position). 称这两种头寸互为**相反头寸** (opposite position).

本书在该部分的中心主题就是寻找一种用来确定衍生品初始价格的方法, 其中衍生品的初始价格为标的资产价格的函数. 这就是**衍生品定价问题**.

衍生品定价问题相对容易解决的唯一时刻就是衍生品的到期时刻. 例如, 如果某一衍生品给你一个权力, 在到期时可以以 \$100 每股的价格购买一股票, 那么, 如果那时的股价低于 \$100, 则该期权没有价值. 另一方面, 如果那时股价为 \$110, 那么衍生品的价值为 \$10. 更一般地, 如果那时股价为 S, 假设不涉及其他额外费用, 则期权价值为 $\max\{S-100,0\}$.

在到期前的任何时候, 衍生品当前价值与标的资产当前价值之间的关系是复杂的, 这就是衍生品定价理论很复杂的原因. 在已知当前情况下, 处理这种复杂关系的唯一方式就是通过假设.

假设

金融市场是复杂的. 与众多复杂系统一样, 建立该系统的数学模型需要做一些简单化假设.

在我们的分析过程中, 将作出一些简单化假设. 例如, 我们假设市场是一个**完美市场**, 意思是该市场满足:

- 没有手续费或交易成本,
- 借贷利率相等,
- 没有卖空限制.

当然, 在现实世界中不存在这样的完美市场, 但是该假设会使分析更简单, 也使我们将精力集中在一些重要问题上, 而这些问题在较少的条件限制下就不那么明显了.

套利

令人惊奇的是, **套利**这个术语包含两种含义. 一般地, 从非技术角度上说, 该术语用来表示一种条件, 在该条件下, 不管环境如何, 投资者能够保证获利.

该术语在技术层面上的使用更广泛, 其含义就有点不同了. **套利机会**是一种投资机会, 保证不会有损失且可能 (存在正的概率) 有收益. 注意不能保证该收益有多大, 但保证不会有损失. 在如何度量收益上我们必须非常小心. 例如, 如果今天的 \$100 在 1 年后变为 \$100.01, 这是真的收益吗? 换句话说, 你愿意参与这样的投资吗? 可能不会, 因为毫无疑问有个无风险收入选择, 比如把钱存在联邦担保的银行里可能会得到更好的收益.

如我们将看到的, 资产定价背后的重要原理就是市场会尽力避免套利. 更具体地说, 如果有一套利机会存在, 那么价格就会被调整以使得该套利机会消失.

举一个简单的例子, 假设在纽约每盎司黄金的价格为 \$380.10, 而伦敦为 \$380.20. 那么投资者可以在纽约购买黄金然后在伦敦卖出, 每盎司获利 10 美分 (假设交易成本没有全部吞没 10 美分). 然而, 在纽约市场上购买黄金会使其在纽约的价格上涨, 而在伦敦市场上出售黄金会使其在伦敦的价格下跌, 结果就是不再有套利.

无套利原理

关于定价的无套利原理其实很简单. 假设有两个投资组合 (包含股票、债券、衍生品等资产), 我们将这两个投资组合标记为组合 A 和组合 B. 考虑两个时期: 初始时期 $t=0$ 和未来时期 $t=T$.

从而每个投资组合有初始价值 (在 0 时的价值) 和最终价值 (在 T 时的价值) 或损益. 记两个组合的初始价值分别为 $\mathcal{V}_{A,0}$ 和 $\mathcal{V}_{B,0}$, 最终价值为 $\mathcal{V}_{A,T}$ 和 $\mathcal{V}_{B,T}$. 组

合 A 的价值如图 1 所示, 对组合 B 也有类似的图.

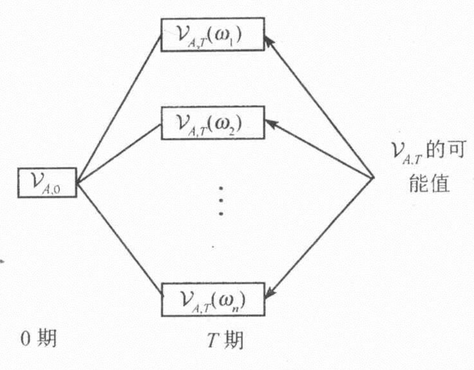


图 1　组合 A 的价值

从图中可以看到, 组合 A 的初始价值已知或者能被确定. 另一方面, 组合 A 的最终价值在 $t=0$ 时未知. 事实上, 我们假设该值依赖 T 时的经济状态, T 时有 n 个可能状态 $\omega_1, \cdots, \omega_n$. 因此, 最终价值 $\mathcal{V}_{A,T}$ 实际上是这些状态的一个函数.

同样地, 我们假设组合 B 的初始价值已知或者能被确定, 最终价值为经济状态的一个函数.

现在, 考虑在任何经济状态下组合 A 和组合 B 都有一模一样的损益时会发生什么, 即对 $i=1, \cdots, n$, 有

$$\mathcal{V}_{A,T}(\omega_i) = \mathcal{V}_{B,T}(\omega_i).$$

则无套利原理暗示着两个组合的初始价值一定相等, 即

$$\mathcal{V}_{A,0} = \mathcal{V}_{B,0}.$$

假设有 $\mathcal{V}_{A,0} > \mathcal{V}_{B,0}$, 那么在完美市场假设下, 投资者可以买入较便宜的组合 B 而卖出较贵的组合 A, 将价差据为己有. 到 T 时, 不管经济处于什么状态, 那时投资者的收入与支出相等. 因此, 他最后没什么损失而享受着最初的利润. 这就是套利.

从而我们看到无套利原理可以用来定价投资组合, 也就是说用来确定投资组合的初始价值. 例如, 为了定价投资组合 A, 我们所需要做的所有事情就是寻找另一投资组合 B, 该组合与组合 A 有相同的损益函数且其初始价值已知. 故组合 A 的初始价值一定等于组合 B 的初始价值.

可以以其他形式使用无套利原理来确定价格. 例如, 如果两个组合的初始价值相等, 那么不可能有一个组合的收益总高于另一组合的收益.

在本书中我们会看到许多利用无套利原理的例子.

第 1 章　概率论 1：离散概率引论

资产定价涉及对未来事件的预测，同样也很依赖概率论知识. 在这一章中，我们开始讨论概率论的基础知识. 考虑到本书后面涉及的主题对概率论知识的需要，我们在后面的章节里还会继续讨论.

概率论似乎起源于对赌博结果的预测，且一般认为它成为一门正式学科是始于 1654 年夏天 Blaise Pascal 和 Pierre de Fermat 两位数学家的一系列书信往来.

1.1　综　　述

在概率论的研究中，典型的场景是试验，比如说掷一对骰子、给病人开药或者预测股票的未来价格. 那么关键点就是试验必须有一个容易确定的可能结果集，该集合被称作试验的样本空间.

样本空间的子集，也就是可能结果的子集，被称作事件. 当特定事件里面的一个结果发生时，我们就说该事件发生了. 因此，比如说在掷骰子的试验中得到的点数和为 7，病人吃药后温度下降到华氏 98.6°，或者股票的价格上升了 10%，这些都可以看成事件.

接下来，必须确定一种测量试验中各种各样的事件发生的概率或者可能性的方法. 更加特别的是，一个事件发生的概率是介于 0 到 1 之间的实数，该数测量了试验发生的结果在该事件中的可能性. 概率为 0 表示该事件不可能发生，而概率为 1 表示该事件一定会发生.

用来确定这些概率的方法并不是概率论这门学科主要研究的. 一般说来有两种方法，一种就是简单地假定概率. 比如，我们考虑抛一枚硬币的试验，那么假设该硬币是均匀的就等价于假定硬币出现正面和反面的概率都是 1/2. 另外一种自然的方法就是统计，利用经验数据来确定概率. 比方说，如果抛 10000 次，出现正面的次数 5003 次，则我们可以说出现正面的概率等于 5003/10000.

概率论的有趣很显然依赖于样本空间的本性. 当处理有限的样本空间时，概率论的基本概念所要求的数学机理非常少，因为在这种情形下，对样本空间中的每一个结果都能够很好地赋予概率，就像上面抛硬币这个例子. 我们很快就将看到，全部的要求就是所有概率值在 0 到 1 之间且加起来等于 1. 那么一个事件的概率就等于该事件所包含的结果的概率和. 有限概率论通常是指有限样本空间的概率论.

举例说明，假设根据对市场的研究，对当前每股价格为 \$100 的股票，我们确定

该股票当天收盘价可能为 \$99、\$100 或者 \$101. 因此, 我们得到一个试验, 其样本空间由该股票所有可能的价格组成

$$\Omega=\{99,100,101\}.$$

更进一步, 通过对该股票历史价格的研究, 我们可以确定经验概率如下

$$\mathbb{P}(99)=0.25,\quad \mathbb{P}(100)=0.5,\quad \mathbb{P}(101)=0.25.$$

在该例中, 若收盘时股票价格没有下降, 即 $\{100,101\}$ 这个事件中的概率为 $\mathbb{P}(100)+\mathbb{P}(101)=0.75$.

对样本空间的状态无限可数的概率论也可以做同样研究, 至少在起始阶段. 其次, 概率值能够赋予给样本空间中的单个结果. 然而, 这涉及无穷和收敛问题. 离散概率这个术语一般是指样本空间有限或无限可数的概率. 在离散概率方面已经有了许多专著.

作为一个离散样本空间 (不一定有限) 的例子, 我们考虑直到首次出现正面才结束的这样一个抛硬币的试验. 结果就是首次出现正面时所需要抛硬币的次数. 开始我们不能确定试验结果集合是任何有限样本空间, 因为事先没有办法知道在首次出现正面之前需要抛多少次. 所以, 样本空间一定是如下集合

$$\Omega=\{1,2,3,\cdots\}$$

由所有的正整数组成. 确实, 有人会说最终肯定会出现正面, 因为如果不是, 那么该集合也就不能表示所有可能的结果了.

如果该硬币是均匀的, 也就是说出现正面和反面的可能性相等, 那么首次出现正面的**等待时间**等于 k 的概率为

$$\mathbb{P}(\text{第 } k \text{ 次才出现正面})=\frac{1}{2^k}.$$

因为无限和

$$\sum_{k\geqslant 1}\frac{1}{2^k}$$

收敛到 1, 我们得到一个合理的概率测度 (我们将会精确地定义该术语).

为了解释如上概率值为什么是合理的, 显然当 $k=1$ 时首次出现正面的概率为 $1/2=1/2^1$. 当 $k=2$ 时首次出现正面的唯一情形就是第一次出现反面且第二次出现正面. 但是对于这两次抛硬币会出现四种等可能性的结果

$$(H,H),\quad (H,T),\quad (T,H),\quad (T,T).$$

因此首次出现正面的合理概率为 $1/4 = 1/2^2$. 根据同样的理由可以推广到更大的数值 k.

我们不想让读者感觉到离散概率比样本空间不可数的非离散概率来得容易. 事实也并非如此. 然而, 对离散概率的基本理解所需要的数学背景弱很多却是事实. 例如, 与有限概率一般不要求了解极限概念一样, 离散概率一般不要求了解积分概念.

对非离散情形, 令人关注的一点是需要更复杂的数学. 比如, 想像一个快要 (或者已经宣布) 破产的公司的股票. 该公司股票价格变为 0 只是时间上的问题. 我们就称它为股票的失效时间 (time to failure). 至少在理论上, 该事件的等待时间可以为任何正实数 (假设该股票一天 24 小时都可交易), 因此样本空间 Ω 由所有的正实数组成, 这是不可数的.

与离散样本空间的情形不一样, 我们不能简单地给每一个失效时间赋予一个概率值, 这是因为无限个不可数的正数的和不可能有限, 更不用说等于 1 了. 与其试着给每一个结果确定概率值, 不如直接给事件赋予概率. 然而, 并不是把样本空间的所有子集都当成事件看待. 这个问题牵涉到更深的知识, 我们就不在此讨论了.

给事件赋予概率最直接直观的方法就是使用函数图像. 见图 1.

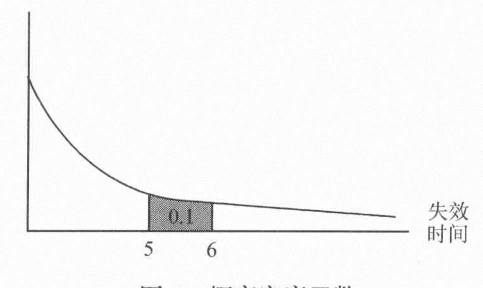


图 1　概率密度函数

图 1 显示了一个能确定任意时间段内失效的概率值的函数的图像. 特别地, 曲线下面的面积就指定了概率值. 例如, 在第 5 天到第 6 天之间失效的概率就等于曲线下面夹在垂线 $x = 5$ 和 $x = 6$ 之间的面积, 为 0.1. 该函数叫做概率密度函数. 概率密度函数, 比如被教授经常用来测定学生们等级的著名钟型曲线, 经常但不总是用来确定非离散情形下的概率. 确实, 有些概率测度是不能用一个概率密度函数来确定的.

在任何情形下, 我们希望大家明白的是非离散情况比离散情况要求更多的数学机理, 即使是概率值的具体化.

对离散时间定价模型的研究我们只需用到有限概率. 在该书的较后面部分我们还将讨论一般理论 (包括非离散情形) 中的一些方面, 作为讨论布莱克 – 舒尔斯期权定价公式的准备.

因此让我们一步一步地来熟悉有限概率方面的基本原理. 毕竟这不是一本有关概率论的课本, 所以我们尽量简明扼要地陈述需要用到的知识. 在以后的章节, 我们还将进一步拓宽对离散概率的讨论, 包括学习离散时间定价模型所必须掌握的内容, 这有助于我们更深刻地理解一般离散时间定价模型.

1.2 概 率 空 间

我们从最主要的定义开始.

定义 1 *有限概率空间是指二元对 $(\Omega, \mathbb{P})$, 其中 Ω 为有限非空集合, 称为样本空间; $\mathbb{P}$ 是定义在 Ω 的所有子集上的实值函数, 称为 Ω 上的概率测度. 函数 $\mathbb{P}$ 必须满足如下性质:*

1) (*值域*) *对所有的 $A \subseteq \Omega$,*

$$0 \leqslant \mathbb{P}(A) \leqslant 1.$$

2) (*Ω的概率*)

$$\mathbb{P}(\Omega) = 1.$$

3) (*可加性*) *如果 A 和 B 不相交, 则有*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

在上下文中, Ω 的子集被称作事件.

如先前提及的, 样本空间用来表示由一个试验所有可能产生的结果组成的集合. 一个特定结果 ω 的概率 $\mathbb{P}(\omega)$ 用来表示试验结果为 ω 的可能性.

另一方面, 这是显而易见的. 正式地说, 我们所关心的一切就是 Ω 为有限非空集合和 $\mathbb{P}$ 为满足如上性质的概率测度. 性质 2) 说的是, 由整个样本空间形成的事件是必然事件, 也就是说任何试验结果肯定会在样本空间里面. 性质 3) 说的是如果两个事件没有公共部分, 那么其中至少有一个事件发生的可能性为两个事件发生可能性之和. 注意这里两个事件不相交非常重要.

有时我们会误用一个概念, 就是把 Ω 本身当成一个概率空间. 在这种情况下, 概率测度 $\mathbb{P}$ 仍然存在. 只是我们没有严格地提及它而已. 学生们应该尽量避免这种小过失.

概率质量函数

如果 Ω 是一有限集, 那么对每一个 Ω 里面的元素 ω, 事件 $\{\omega\}$ 就叫做基本事件 (elementary event). 在有限样本空间 Ω 上定义一个概率测度最简单的办法

就是确定所有基本事件的概率. 这等价于我们赋予每个 $\omega \in \Omega$ 一个数值 p_ω, 满足 $0 \leqslant p_\omega \leqslant 1$ 和

$$\sum_{\omega \in \Omega} p_\omega = 1.$$

然后我们就可以按如下方式定义一个概率测度 $\mathbb{P}$,

$$\mathbb{P}(\{\omega\}) = p_\omega.$$

再利用有限可加性将 $\mathbb{P}$ 扩展到所有事件上. 这是一种非常有趣的方法, 任何事件 E 的概率等于 E 所包含的所有基本事件的概率和.

称集合 $\{p_\omega | \omega \in \Omega\}$ 为概率分布, 称如下定义的函数 $f : \Omega \longrightarrow R$,

$$f(\omega) = p_\omega,$$

为概率质量函数 (注意不要混淆概率分布和分布函数这两个概念, 后者有不同的含义, 我们会在后面的章节里给出其定义).

注意到, 概率测度 $\mathbb{P}$ 和概率质量函数 f 之间细微但重要的区别就是 $\mathbb{P}$ 定义在 Ω 的所有子集上而 f 定义在 Ω 的所有元素上.

当一个概率分布给定时, 任何事件 $A \subseteq \Omega$ 的概率为该事件里所有结果的概率和, 也就是

$$\mathbb{P}(A) = \sum_{\omega \in A} p_\omega.$$

进一步地, 如果样本空间里的每个结果有相同的概率, 即为 $1/|\Omega|$, 因此任何事件 E 的概率就简单地等于 E 的大小除以样本空间 Ω 的大小, 即

$$\mathbb{P}(E) = \frac{|E|}{|\Omega|}.$$

例 1 根据某一股票过去几年历史价格的研究表明, 在一月份里其价格将达到最大值的概率如下

$$\begin{aligned}
\mathbb{P}(0-4.99) &= 0.65,\\
\mathbb{P}(5-9.99) &= 0.2,\\
\mathbb{P}(10-14.99) &= 0.1,\\
\mathbb{P}(15-19.99) &= 0.04,\\
\mathbb{P}(20-24.99) &= 0.01.
\end{aligned}$$

那么在该月里股票价格能够达到 \$10 的概率是多少? 股票价格要么没有达到 \$5 要么达到了 \$20 的概率又是多少呢?

解　在该月股票价格能达到 \$10 当且仅当在该月里股票价格最大值至少是 \$10. 因此

$$\begin{aligned}\mathbb{P}(\text{价格达到}10) &= \mathbb{P}(10-14.99)+\mathbb{P}(15-19.99)+\mathbb{P}(20-24.99)\\ &= 0.1+0.04+0.01=0.15.\end{aligned}$$

同理

$$\begin{aligned}\mathbb{P}(\text{价格没达到}5\text{或者达到了}20) &= \mathbb{P}(0-4.99)+\mathbb{P}(20-24.99)\\ &= 0.65+0.01=0.66.\end{aligned}$$

概率论有自己的专业词汇术语, 即使就像集合不相交这样简单的概念.

定义 2　当两个事件 A 和 B 作为集合不相交时, 我们就说它们相互排斥. 当一系列事件组成的群集 $\{A_1,\cdots,A_n\}$ 满足

$$A_i\cap A_j=\varnothing$$

对所有不同的 i,j. 我们就说该群集两两相互排斥.

由概率空间的定义可得到下面一些简单结果.

定理 1　设 (Ω,P) 是一有限概率空间. 则

1) (空集的概率)

$$\mathbb{P}(\varnothing)=0.$$

2) (单调性)

$$A\subseteq B\Rightarrow \mathbb{P}(A)\leqslant \mathbb{P}(B).$$

3) (补集的概率)

$$\mathbb{P}(A^c)=1-\mathbb{P}(A).$$

4) (有限可加性)　如果 $\{A_1,\cdots,A_n\}$ 是 Ω 中有限两两相互排斥的群集, 则

$$\mathbb{P}(A_1\cup\cdots\cup A_n)=\mathbb{P}(A_1)+\cdots+\mathbb{P}(A_n).$$

划分和全概率定理

下面的简单概念将在我们要讨论的衍生品定价模型中发挥重要作用.

定义 3 设 Ω 是一非空集合. 则 Ω 的一个划分是一群集 $\mathcal{P}=\{B_1,\cdots,B_n\}$, Ω 的子集 $B_1,\cdots,B_n$ 称为划分的块, 有如下性质:

1) 这些块两两不相交

$$B_i \cap B_j = \varnothing$$

对所有的 i,j.

2) 块的并为 Ω, 即

$$B_1 \cup \cdots \cup B_n = \Omega.$$

下面这个重要定理告诉我们, 如果能确定 E 与一个划分的每个块相交部分的概率, 那么我们就能确定事件 E 的概率. 其证明留个读者.

定理 2 (全概率公式) 设 $(\Omega,\mathbb{P})$ 是一样本空间, 事件 $E_1,\cdots,E_n$ 形成 Ω 的一个划分. 则对 Ω 中的任何事件 A, 有

$$\mathbb{P}(A)=\sum_{k=1}^{n}\mathbb{P}(A\cap E_k).$$

1.3 独 立 性

一枚均匀的硬币就是其出现正面的概率为 1/2. 确实, 这就是均匀硬币的定义. 假想我们抛一枚均匀的硬币 99 次且每次都出现正面, 诚然这是不太可能的事情, 但是仍然有可能发生. 你愿意打赌说第 100 次抛后的结果还是正面? 很多人都不会, 原因是既然正面已经连续出现了这么多次, 那么反面正在"迟到的"路上.

然而事实上是, 每次抛硬币出现的结果与其他几次抛的结果独立, 因此第 100 次出现正面的概率仍是 1/2, 而不管以前的结果如何.

在这一点上产生混淆的原因可能是牵涉到先前连续出现了 99 次正面, 当然这样的概率非常小. 但是一旦发生了, 这样极端不太可能性影响就要剔除, 我们重新回到考虑抛一硬币出现各种结果的可能性.

直觉告诉我们, 如果一个事件发生并不影响另外一个事件发生的概率, 则这两个事件独立. 当在后面讨论条件概率的时候, 我们将给出该概念的精确定义. 我们现在正式定义独立性.

定义 4 称概率空间 $(\Omega,\mathbb{P})$ 上的事件 E 和 F 独立, 如果两事件都发生的概率等于各个事件发生的概率乘积, 用记号表示为

$$\mathbb{P}(E\cap F)=\mathbb{P}(E)\mathbb{P}(F).$$

例如，假设某一股票在一天里的价格可能上升也可能下跌，某一债券也是. 如果我们假定股票与债券的行为独立，则

$$\mathbb{P}(\text{股票上升，债券下跌}) = \mathbb{P}(\text{股票上升})\mathbb{P}(\text{债券下跌}).$$

我们也能定义一个群集的独立性.

定义 5　*称群集 $\{E_1, \cdots, E_n\}$ 独立，如果对任意子群集 $\{E_{i_1}, \cdots, E_{i_k}\}$，有*

$$\mathbb{P}(E_{i_1} \cap \cdots \cap E_{i_k}) = \mathbb{P}(E_{i_1}) \cdots \mathbb{P}(E_{i_k}).$$

注意到当判断三个事件 A, B, C 是否独立时，我们必须检验下面三个条件：

A 和 B 独立，

A 和 C 独立，

B 和 C 独立.

一般地，要判断含有 k 个事件的群集的独立性，我们必须检验 C_k^2 个条件. 因此随着事件数的增加，检验条件数急剧增加.

1.4　二项式概率

在有意义的试验中最简单的一类就是试验结果只有两种的试验. 这样的试验称为伯努利 (Bernoulli) 试验. 这两个结果通常用成功和失败来表示，且成功的概率一般用 p 来表示，因此失败的概率就为 $1-p$.

例如，抛一枚硬币就是伯努利试验，在该试验中，我们可以认为出现正面表示成功，出现反面表示失败 (或者反之). 再举一个与此相关的例子，在我们考虑的一个衍生品定价模型中，股价在任意给定的时刻 t_k 可能从先前的价格 S 上涨到 Su，或者下跌到 Sd，这里 $0 \leqslant d \leqslant 1 \leqslant u$. 因此，在每个时刻 t_k 我们得到一个伯努利试验.

如果一个有成功概率为 p 的伯努利试验被重复 n 次，则被叫做二项式试验. 注意到既然相同的试验被重复，则每次试验的结果独立，也就是说，第 k 次的试验结果不影响第 $k+1$ 次试验的结果. 二项式试验的参数为 p 和 n.

比如，重复抛一硬币 n 次就是二项式试验. 有放回的从一副扑克里面抽取 n 次，每次抽取一张牌，抽取到 A 表示成功，其概率为 p，这是二项式试验. 必要的是我们必须每次重复相同的二项式试验.

因为二项式试验中的单个伯努利试验是独立的，所以很容易计算二项式试验中的任意特定结果的概率，下面的例子说明了这点. 确实，我们将在以后更加仔细地讨论下面例子的一般情形.

例 2 考虑一只股票, 其价格会在如下 6 个时刻的任一时刻发生变化

$$t_0 < t_1 < \cdots < t_5.$$

假设股票在 t_0 时的初始价格为 S. 更进一步假设在时间 $[t_k, t_{k+1}]$ 内股价或者上涨到原来的 u 倍, 或者下跌到原来的 d 倍, 其中 $0 < d < 1 < u$, 且股价变化与先前独立. 股价上涨的概率为 p. 因此, 在每个时间段内, 可以看作具有成功概率为 p 的伯努利试验. 而且整个历史价格过程是一个具有参数 p 和 $n = 5$ 的二项式试验.

该二项式试验的典型结果可以记为由 U 和 D 排成长度为 5 的序列, 因此样本空间就是如下集合

$$\Omega = \{U, D\}^5.$$

例如, 序列 $UUDUD$ 表示在时间段 $[t_0, t_1]$, $[t_1, t_2]$ 和 $[t_3, t_4]$ 内股价上升, 而在 $[t_2, t_3]$ 和 $[t_4, t_5]$ 上股价下跌.

为了计算出现该结果的概率, 我们利用单个伯努利试验之间相互独立的关系, 得到它们相交的概率等于各自的概率乘积. 故有

$$\mathbb{P}(UUDUD) = pp(1-p)p(1-p) = p^3(1-p)^2.$$

显然, 每个元素 $\omega \in \Omega$ 的概率只依赖 ω 中 U 和 D 的数目, 而与 U 和 D 的排列顺序无关. 因此, 如果令

$$N_U(\omega) = \omega\text{中 } U \text{ 的数目},$$
$$N_D(\omega) = \omega\text{中 } D \text{ 的数目}.$$

则

$$\mathbb{P}(\omega) = p^{N_U(\omega)}(1-p)^{N_D(\omega)}.$$

我们来计算股价刚好有 3 次上涨的概率. 乏味的方法就是列举出满足条件的所有可能的价格过程. 如下表.

$UUUDD$	$UUDUD$
$UUDDU$	$UDUUD$
$UDUDU$	$UDDUU$
$DUUUD$	$DUUDU$
$DUDUU$	$DDUUU$

如上表, 有 10 种情形满足条件, 且每种情形的概率为 $p^3(1-p)^2$, 因此所求概率为 $10p^3(1-p)^2$.

现在非常容易使该结果一般化. 刚好有 k 次上涨（$n-k$ 次下跌）的概率正是

$$C_n^k p^k (1-p)^{n-k}.$$

我们已经建立了下面的有用结论.

定理 3 考虑参数为 p 和 n 的二项式试验. 该试验的样本空间为 $\Omega = \{s, f\}^n$, 其中 s 代表成功, f 代表失败. 对任何 $\omega \in \Omega$, 令

$$N_s(\omega) = \omega\text{中}s\text{的数目}.$$

1) 如果 $\omega \in \Omega$, 则

$$\mathbb{P}(\omega) = p^{N_s(\omega)}(1-p)^{n-N_s(\omega)}.$$

2) 恰好有 k 次成功的概率为

$$\mathbb{P}(\text{恰好成功 } k \text{ 次}) = C_n^k p^k (1-p)^{n-k}.$$

例 3 从一副扑克牌里面有放回地抽取四张牌, 请问至少抽到 3 张 A 的概率为多少?

解 至少抽到 3 张 A 的概率等于刚好抽到 3 张 A 的概率和刚好抽到 4 张 A 的概率之和. 因为我们进行的是二项式试验, 每次成功得到一张 A 的概率为 $p = 4/52 = 1/13$, 所以有

$$\begin{aligned}
\mathbb{P}(\text{至少抽到3张}A) &= \mathbb{P}(\text{刚好抽到3张}A) + \mathbb{P}(\text{刚好抽到4张}A) \\
&= C_4^3\left(\frac{1}{13}\right)^3\left(\frac{12}{13}\right)^1 + C_4^4\left(\frac{1}{13}\right)^4\left(\frac{12}{13}\right)^0 \\
&= \frac{49}{28561} \\
&\approx 0.0017.
\end{aligned}$$

这是非常小的一个概率.

定理 3 和例 3 中描述的概率分布非常重要.

定义 6 令 $0 < p < 1$, n 是正整数, $\Omega = \{0, \cdots, n\}$. Ω 上质量函数为

$$b(k; n, p) = C_n^k p^k (1-p)^{n-k}, \quad k = 0, \cdots, n$$

的概率分布称为二项分布. 该分布给出了在参数为 p 和 n 的二项式试验中刚好有 k 次成功的概率.

图 2 给出了两个二项分布.

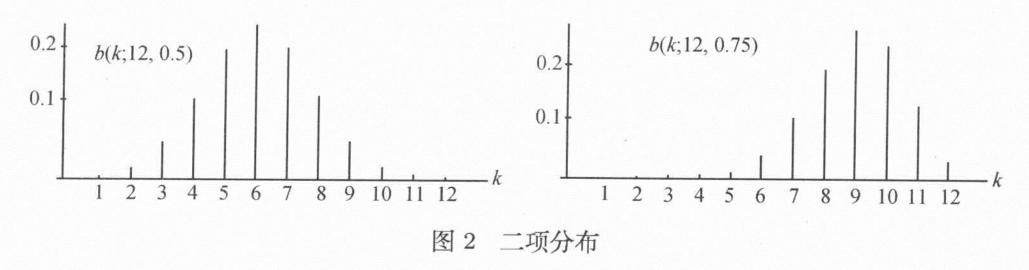


图 2 二项分布

经验概率和理论概率

前面已经间接提到, 有两种共同的方法确定概率. 考虑在例 2 中确定概率 p 的值这个问题. 这是股价上涨的概率.

一种方法就是细心检查股价在一段时间内的历史价格情况. 然后我们就能用股价上涨的次数除以总时刻数来得到 p 的值. 例如, 如果在过去 10000 个时刻点股价上涨的次数为 5003 次, 那么我们就能令

$$p = \frac{5003}{10000}. \tag{1.4.1}$$

当然, 很自如得到下跌的概率为

$$1 - p = 1 - \frac{5003}{10000} = \frac{4997}{10000}. \tag{1.4.2}$$

因为这些概率值是通过分析经验数据而得到的结果, 或者说至少是通过对实际物理现象分析而得到的结果, 它们被称作经验概率 (empirical probabilities).

另一方面, 可能从头到尾都缺少用来分析的任何实际数据, 我们可以简单地假定 $p = 1/2$. 这类概率叫做理论概率. 就像我们将看到的一样, 这两类概率在金融数学中都有它们自己的地位.

1.5 随机变量

下面的定义非常重要.

定义 7 称一个定义在有限样本空间 Ω 上的实值函数 $X : \Omega \longrightarrow R$ 为 Ω 上的随机变量. Ω 上的所有随机变量形成的集合记为 $RV(\Omega)$.

就像定义所叙述的, 对于有限 (或者离散) 的概率空间来说, 随机变量和实值函数没有什么区别. 但是, 我们在后面部分将看到对于非离散的样本空间来说, 不是所有的实值函数都能够用来表示随机变量.

因为 $RV(\Omega)$ 正是 Ω 上所有函数的集合，那么在函数的相加和点乘法则下，$RV(\Omega)$ 是一个向量空间. 因此，如果 X 和 Y 是 Ω 上的随机变量，$a, b \in R$，则

$$aX + bY$$

也是 Ω 上的随机变量. 同时也注意，两个 Ω 上的随机变量的乘积也是 Ω 上的随机变量.

随机变量中最有用的一类就是那些等同于一些特定事件的随机变量.

定义 8 设 A 是 Ω 中的一个事件. 函数 1_A^Ω 定义如下

$$1_A^\Omega = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

称它为 A 的示性函数（或者示性随机变量）. 当 Ω 不混淆时，我们也可以将 A 的示性函数写成 1_A.

例 4 令

$$\Omega = \{0.5, 0.75, 1, 1.25, 1.5, 1.75\}$$

为可能的联邦折现率的样本空间. 考虑一个公司，它的股票价格随利率波动. 股价可以表示成 Ω 上的随机变量 S. 例如

$$\begin{aligned} S(0.5) &= 105, \\ S(0.75) &= 100, \\ S(1) &= 100, \\ S(1.25) &= 100, \\ S(1.5) &= 95, \\ S(1.75) &= 90. \end{aligned}$$

$\{S = 100\}$ 事件和由折现率 $\{0.75, 1, 1.25\}$ 所形成的事件是相同的，即

$$\{S = 100\} = \{0.75, 1, 1.25\}.$$

例 5 考虑掷两个均匀骰子并记录下每个骰子的点数的试验. 样本空间由 36 对有序数对组成

$$\Omega = \{(1,1), (1,2), (1,3), \cdots, (6,4), (6,5), (6,6)\}.$$

既然骰子是均匀的，每个有序数对发生的机会一样，因此每个结果发生的概率均为 $1/36$.

然而，在一些赌博中，我们感兴趣的只是两个点数的和. 因此，我们定义随机变量 $S:\Omega \longrightarrow R$,

$$S(a,b) = a + b.$$

得到 7 点和的事件 $\{S = 7\}$ 为

$$\{S=7\} = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}.$$

有

$$\mathbb{P}(\text{点数和为}7) = \mathbb{P}(S=7) = \frac{6}{36} = \frac{1}{6}.$$

可能关于随机变量最基本的实质就是用它们来表示事件. 事实上，很多时候我们并不关心 S 的真实值，而只关心由这些值表示的事件. 例如，在前面的例子里如果我们用“倍和”随机变量

$$D((a,b)) = 2(a+b).$$

那么 D 描述了同样的事件. 比如，$\{S=7\} = \{D=14\}$.

当然也不全是这样的情况. 有些随机变量的真实值是非常重要的. 例如股价随机变量、利率随机变量、成本随机变量. 所有的随机变量都是用来描述我们感兴趣的一些特定事件集合.

随机变量的概率分布

设 X 是概率空间 $(\Omega,\mathbb{P})$ 上的随机变量，其中 $\Omega = \{\omega_1,\cdots,\omega_n\}$. 因 Ω 有限，故 X 取有限个可能值，记为 $\mathcal{A} = \{x_1,\cdots,x_m\}$. 对每个 x_i，我们能定义事件

$$\{X = x_i\} = X^{-1}(x_i) = \{\omega_j | X(\omega_j) = x_i\}.$$

这正好是 x_i 的原象. 表达式 $\{X = x_i\}$ 是随机变量刻画的事件的最普遍的符号. 因为随机变量的值域是一实数集合，我们也能够定义事件如下

$$\{X \leqslant x_i\} = X^{-1}((-\infty, x_i)) = \{\omega_j | X(\omega_j) \leqslant x_i\}.$$

事件

$$\{X = x_1\}, \{X = x_2\}, \cdots, \{X = x_m\}$$

形成了样本空间 Ω 的一个划分，也就是说，这些事件两两不相交且它们的并为 Ω (因为 X 必须定义在 Ω 上的所有元素). 因此,

$$\sum_{i=1}^{m}\mathbb{P}(X=x_i)=1.$$

注意习惯性用较为简单的记号 $\mathbb{P}(X=x)$ 来代替 $\mathbb{P}(\{X=x\})$.

很自然, $\mathbb{P}(X=x_i)$ 形成了集合 $\mathcal{A}=\{x_1,\cdots,x_m\}$ 上的一个概率分布，其中, $\mathcal{A}=\{x_1,\cdots,x_m\}$ 为 $\mathbb{R}$ 的一个子集. 因此，随机变量 X 描述了 $\mathcal{A}$ 上的一个概率测度

$$\mathbb{P}_X(\{x_i\})=\mathbb{P}(X=x_i)$$

称为由 X 诱导的概率测度 (或者概率分布). 相应的概率质量函数 $f:\mathcal{A}\longrightarrow\mathbb{R}$ 为

$$f(x_i)=\mathbb{P}(X=x_i),$$

称为 X 的概率质量函数.

因此，比方说随机变量 X 具有参数为 p 和 n 的二项式分布等同于说 X 的取值为 $\{0,1,\cdots,n\}$ 且 X 的概率质量函数为

$$\mathbb{P}(X=k)=b(k;n,p)=C_n^k p^k(1-p)^{n-k}.$$

这种情况下一般也说 X 是二项式分布.

有关随机变量的这些事实如此重要，所以忍不住再次重复一下. 随机变量被用来等同样本空间的某些相关事件. 而且随机变量能转换概率测度，把它等同的样本空间中的事件转变到随机变量在 $\mathbb{R}$ 中的值域.

定义 9 称从样本空间 Ω 到向量空间 $\mathbb{R}^n$ 的一个函数 $X:\Omega\longrightarrow\mathbb{R}^n$ 为 Ω 上的随机向量.

样本空间 Ω 上的所有随机向量形成的集合 $RV^n(\Omega)$ 在函数之间的加法和点乘法则为向量空间.

例 6 令

$$\Omega=\{0.5,,0.75,1,1.25,1.5,1.75\}$$

为可能的联邦折现率的样本空间. 考虑一个公司，它的股票价格随利率波动. 当然债券价格也随利率波动. 我们定义价格随机向量 $S:\Omega\longrightarrow\mathbb{R}^2, S(\omega)=(s,b)$, 其中 s 和 b 分别是折现率为 ω 时的股票价格和债券价格. 例如,

$$S(0.5)=(105,112),$$

意思是当折现率为 0.5% 时，股票价格为 105，债券价格为 112.

随机变量的独立性

称样本空间 Ω 上的两个随机变量 X 和 Y 独立, 如果对任意的 x, y, 事件 $\{X = x\}$ 和 $\{Y = y\}$ 都独立. 直观上说的是, 要知道一个随机变量的值与是否知道另外一个随机变量的值无关. 下面有更正式的定义, 在这里用 $\mathbb{P}(X = x, Y = y)$ 简记 $\mathbb{P}(\{X = x\} \cap \{Y = y\})$.

定义 10 称 Ω 上的两个随机变量 X 和 Y 独立, 如果对所有的 $x, y \in \mathbb{R}$, 有

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

更一般地, 称随机变量 $X_1, \cdots, X_n$ 是独立的, 如果对所有的 $x_1, \cdots, x_n \in \mathbb{R}$, 有

$$\mathbb{P}(X_1 = x_1, \cdots, X_n = x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i).$$

1.6 期　望

期望这个概念在金融数学中有着重要的地位.

定义 11 设 X 是有限概率空间 $(\Omega, \mathbb{P})$ 上的一随机变量, 其中 $\Omega = \{\omega_1, \cdots, \omega_n\}$. X 的期望 (或者均值) 定义如下

$$\mathcal{E}_{\mathbb{P}}(X) = \sum_{i=1}^{n} X(\omega_i)\mathbb{P}(\omega_i).$$

求和中的项的形式是: X 在 ω 上的值乘以 ω 发生的概率. 如果 X 只取几个不同的值 $\{x_1, \cdots, x_m\}$, 则也有

$$\mathcal{E}_{\mathbb{P}}(X) = \sum_{i=1}^{m} \mathbb{P}(X = x_i).$$

注意两个求和式中的上限不同. 这是 X 的加权平均和, 其权重为相应发生的概率大小. X 的期望也记作 μ_X.

期望函数 $\mathcal{E} : RV(\Omega) \longrightarrow \mathbb{R}$ 把 $(\Omega, \mathbb{P})$ 上的随机变量映射到实数. 这类函数的一个重要性质就是它是线性的.

定理 4 期望函数 $\mathcal{E} : RV(\Omega) \longrightarrow \mathbb{R}$ 是线性的, 即对任意随机变量 X, Y 和实数 a, b, 有

$$\mathcal{E}(aX + bY) = a\mathcal{E}(X) + b\mathcal{E}(Y).$$

证明　假设 X 取值为 $\{x_1, \cdots, x_n\}$, Y 的取值为 $\{y_1, \cdots, y_m\}$. 则 $aX+bY$ 取的值为 $ax_i + by_j, i = 1, \cdots, n, j = 1, \cdots, m$. 为了计算 $aX+bY$ 的期望, 我们考虑如下事件

$$E_{i,j} = \{X = x_i, Y = y_j\},$$

其中 $i = 1, \cdots, n, j = 1, \cdots, m$. 这些事件形成了 Ω 的一个划分, 且 $aX+bY$ 在 $E^{i,j}$ 上取常值 $ax_i + by_j$, 利用全概率公式有

$$\begin{aligned}
\mathcal{E}(aX+bY) &= \sum_{i=1}^{n}\sum_{j=1}^{m}(ax_i + by_j)\mathbb{P}(X = x_i, Y = y_j) \\
&= a\sum_{i=1}^{n} x_i\Big[\sum_{j=1}^{m}\mathbb{P}(X = x_i, Y = y_j)\Big] \\
&\quad + b\sum_{j=1}^{m} y_j\Big[\sum_{i=1}^{n}\mathbb{P}(X = x_i, Y = y_j)\Big] \\
&= a\sum_{i=1}^{n} x_i\mathbb{P}(X = x_i) + b\sum_{j=1}^{m} y_j\mathbb{P}(Y = y_j) \\
&= a\mathcal{E}(X) + b\mathcal{E}(Y).
\end{aligned}$$

定理得证.

随机变量函数的期望

注意到, 如果 $f : \mathbb{R} \longrightarrow \mathbb{R}$ 是实变量的实值函数, 且 X 是随机变量, 则它们的复合 $f(X) : \Omega \longrightarrow \mathbb{R}$ 也是一个随机变量. 在有限概率空间上, 仅仅用到函数的复合也是函数这个事实.

随机变量 $f(X)$ 的期望等于

$$\mathcal{E}_{\mathbb{P}}(f(X)) = \sum_{i=1}^{n} f(X(\omega_i))\mathbb{P}(\omega_i),$$

或者

$$\mathcal{E}_{\mathbb{P}}(f(X)) = \sum_{i=1}^{n} f(x_i)\mathbb{P}(X = x_i).$$

当没有必要强调概率测度时, 我们将省略 $\mathcal{E}_{\mathbb{P}}$ 的下标而直接写成 $\mathcal{E}$, 但重要的是我们应在心里记着期望依赖该特定的概率测度.

例 7　考虑一只股票, 它的当前价格为 100, 在 T 时刻的价格依赖于经济状况, 可能的经济状况为

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}.$$

各经济状况出现的概率如下

$$\mathbb{P}(\omega_1) = 0.2, \quad \mathbb{P}(\omega_2) = 0.3, \quad \mathbb{P}(\omega_3) = 0.3, \quad \mathbb{P}(\omega_4) = 0.2.$$

股票价格这个随机变量取值情况如下

$$S(\omega_1) = 99, \quad S(\omega_2) = 100, \quad S(\omega_3) = 101, \quad S(\omega_4) = 102.$$

如果我们现在购买一份股票, 则 T 时的期望收入为

$$\mathcal{E}(S) = 99 \times 0.2 + 100 \times 0.3 + 101 \times 0.3 + 102 \times 0.2 = 100.5.$$

因此期望利润为 $100.5 - 100 = 0.5$. 现考虑一衍生产品, 它的收益 D 为股票价格的函数, 定义如下

$$\begin{aligned} D(99) &= -4, \\ D(100) &= 5, \\ D(101) &= 5, \\ D(102) &= -6. \end{aligned}$$

所以 D 是 Ω 上的随机变量. 该衍生产品的期望收益为

$$\begin{aligned} \mathcal{E}(\text{收益}) &= D(99)\mathbb{P}(99) + D(100)\mathbb{P}(100) \\ &\quad + D(101)\mathbb{P}(101) + D(102)\mathbb{P}(102) \\ &= (-4) \times 0.2 + 5 \times 0.3 + 5 \times 0.3 - 6 \times 0.2 \\ &= 1. \end{aligned}$$

前面这个例子指出了期望的一个重要性质. 期望值很少会是值的期望. 在这个例子中, 我们从来没期望能得到 100.5 的收入. 事实上, 该收入是不可能的. 收入必须是样本空间里的一个值. 期望是平均值, 而不是最希望得到的值 (称最希望得到的值为模).

期望和独立

我们已经看到期望算子是线性的, 即

$$\mathcal{E}(aX + bY) = a\mathcal{E}(X) + b\mathcal{E}(Y).$$

很自然地会想到求 $\mathcal{E}(XY)$. 设 X 取值为 $\{x_1,\cdots,x_n\}$, Y 的取值为 $\{y_1,\cdots,y_m\}$. 则乘积 XY 取值为 $x_iy_i, i=1,\cdots,n, j=1,\cdots,m$. 考虑如下事件

$$E_{i,j}=\{X=x_i,Y=y_j\},$$

其中 $i=1,\cdots,n, j=1,\cdots,m$. 这些事件形成了 Ω 的一个划分, 且 $aX+bY$ 在 $E^{i,j}$ 上取常值 ax_i+by_j, 因此

$$\mathcal{E}(XY)=\sum_{i=1}^{n}\sum_{j=1}^{m}x_iy_j\mathbb{P}(X=x_i,Y=y_j).$$

一般地, 我们不能计算 $\mathbb{P}(X=x_i,Y=y_j)$. 但是, 如果 X 与 Y 独立, 则有

$$\begin{aligned}\mathcal{E}(XY)&=\sum_{i=1}^{n}\sum_{j=1}^{m}x_iy_j\mathbb{P}(X=x_i,Y=y_j)\\&=\sum_{i=1}^{n}\sum_{j=1}^{m}x_iy_j\mathbb{P}(X=x_i)\mathbb{P}(Y=y_j)\\&=\Big[\sum_{i=1}^{n}x_i\mathbb{P}(X=x_i)\Big]\Big[\sum_{j=1}^{m}y_j\mathbb{P}(Y=y_j)\Big]\\&=\mathcal{E}(X)\mathcal{E}(Y).\end{aligned}$$

所以我们有个非常重要的定理.

定理 5　如果 X 与 Y 是概率空间 $\{\Omega,\mathbb{P}\}$ 上相互独立的随机变量, 则有

$$\mathcal{E}(XY)=\mathcal{E}(X)\mathcal{E}(Y).$$

该定理能推广到两个以上独立随机变量乘积情形. 例如, 不难知道, 如果 X, Y 和 Z 是独立的, 则 XY 与 Z 也是独立的, 故

$$\mathcal{E}(XYZ)=\mathcal{E}(XY)\mathcal{E}(Z)=\mathcal{E}(X)\mathcal{E}(Y)\mathcal{E}(Z).$$

1.7　方差和标准差

一个随机变量 X 的期望是其分布中心的测量. 而对随机变量取值分散程度的测量通常使用方差和方差的平方根, 即标准差. 标准差的优点体现在它和随机变量有相同的单位. 但是它的缺点是存在平方根.

定义 12　设 X 为随机变量, 期望 μ 有限, 其方差定义为

$$\sigma_X^2=\mathrm{var}(X)=\mathcal{E}((X-\mu)^2).$$

标准差为方差的正平方根, 即

$$\sigma_X = SD(X) = \sqrt{\mathrm{var}(X)}.$$

下一定理给出了方差的一些简单性质.

定理 6 设 X 为随机变量, 期望 μ 有限, 那么

1) $\mathrm{var}(X) = \mathcal{E}(X^2) - \mu^2 = \mathcal{E}(X^2) - \mathcal{E}(X)^2$.

2) 对任意实数 a, 有

$$\mathrm{var}(aX) = a^2\mathrm{var}(X).$$

3) 如果 X 与 Y 是相互独立的随机变量, 则

$$\mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y).$$

4) 如果 c 是常数, 则

$$\mathrm{var}(X+c) = \mathrm{var}(X).$$

我们将证明留作练习.

注意, 方差不像期望算子那样是线性的. 因此

$$\mathrm{var}(aX+bY)$$

和

$$a\mathrm{var}(X) + b\mathrm{var}(Y)$$

是不一样的. 在后面部分还会更进一步讨论该问题.

随机变量的标准化

如果 X 是一个均值为 μ 和方差为 σ^2 的随机变量, 我们能够如下定义一个新的随机变量 Y,

$$Y = \frac{X-\mu}{\sigma}.$$

则有

$$\mathcal{E}(Y) = \mathcal{E}\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}[\mathcal{E}(X) - \mu] = 0$$

和

$$\mathrm{var}(Y) = \mathrm{var}\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}\mathrm{var}(X-\mu) = \frac{1}{\sigma^2}\mathrm{var}(X) = 1.$$

我们看到, Y 的期望为 0, 方差为 1. 称从 X 到 Y 的过程为随机变量 X 的标准化过程.

二项分布的均值

我们能很容易地计算二项分布的均值和方差.

定理 7 设随机变量 X 服从二项分布 $b(k;n,p)$, 那么

$$\mathcal{E}(X)=np,$$
$$\operatorname{var}(X)=np(1-p).$$

证明 令 $q=1-p$. 先考虑期望, 有

$$\begin{aligned}\mathcal{E}(X)&=\sum_{k=0}^{n}k\mathbb{P}(X=k)\\&=\sum_{k=0}^{n}kC_n^k p^k q^{n-k}\\&=np\sum_{k=0}^{n}C_{n-1}^{k-1}p^{k-1}q^{(n-1)-(k-1)}\\&=np\sum_{k=0}^{n-1}C_{n-1}^{k}p^{k}q^{(n-1)-k}\\&=np.\end{aligned}$$

我们将方差的推导留作练习.

1.8 协方差, 相关性和最佳线性估计

我们现在希望挖掘出定义在相同概率空间上的两个随机变量之间的关系.

定义 13 设 X 和 Y 是两个均值有限的随机变量, 定义它们的协方差如下

$$\sigma_{X,Y}=\operatorname{cov}(X,Y)=\mathcal{E}[(X-\mu_X)(Y-\mu_Y)].$$

协方差的一些性质将在下面的定理中给出.

定理 8 协方差满足如下性质：

1) 协方差的期望表示形式

$$\operatorname{cov}(X,Y)=\mathcal{E}(XY)-\mathcal{E}(X)\mathcal{E}(Y).$$

2) (对称性)

$$\operatorname{cov}(X,Y)=\operatorname{cov}(Y,X).$$

3) X 与本身的协方差正好是 X 的方差, 即

$$\operatorname{cov}(X, X) = \sigma_X^2.$$

4) 如果 X 是一个常数随机变量 (就是说 $\sigma_X = 0$) , 则

$$\operatorname{cov}(X, Y) = 0.$$

5) 协方差函数是双线性的, 即

$$\operatorname{cov}(aX + bY, Z) = a\operatorname{cov}(X, Z) + b\operatorname{cov}(Y, Z).$$

6) 协方差的有界性

$$|\operatorname{cov}(X, Y)| \leqslant \sigma_X \sigma_Y.$$

而且取等号时当且仅当 X 或 Y 中的一个是常数或者存在常数 a 和 b, 使得

$$Y = aX + b.$$

证明 我们只证明 6). 如果 X 或 Y 中的一个是常数, 则两边都为 0, 因此假定另外一个条件. 令 t 为实变量, 有

$$\begin{aligned}
0 \leqslant & \mathcal{E}((tX + Y)^2) \\
= & \mathcal{E}(t^2 X^2 + 2tXY + Y^2) \\
= & t^2 \mathcal{E}(X^2) + 2t\mathcal{E}(XY) + \mathcal{E}(Y^2) \\
= & f(t),
\end{aligned}$$

这里 $f(t)$ 是 t 的二次函数. 因为 $f(t)$ 恒大于或等于零且二次项系数为正, 故 $f(t)$ 的判别式非正, 即

$$[2\mathcal{E}(XY)^2] - 4\mathcal{E}(X^2)\mathcal{E}(Y^2) \leqslant 0,$$

或者

$$\mathcal{E}(XY)^2 \leqslant \mathcal{E}(X^2)\mathcal{E}(Y^2).$$

而且, 等号成立 (判别式为零) 当且仅当存在一个 t, 使得 $f(t) = \mathcal{E}((tX + Y)^2) = 0$. 但这又当且仅当 $Y = -tX$(我们假设 Y 不是常数, 这意味着 t 不为 0).

因为这对任何随机变量 X 和 Y 都适应, 我们将它运用到随机变量 $X - \mu_X$ 和 $Y - \mu_Y$ 上得到

$$|\operatorname{cov}(XY)^2| \leqslant \sigma_X^2 \sigma_Y^2,$$

也就是

$$|\mathrm{cov}(X,Y)| \leqslant \sigma_X \sigma_Y,$$

取等号时当且仅当 X 或 Y 至少有一个是常数, 或者存在非零实数 a, 使得

$$Y-\mu_Y = a(X-\mu_X).$$

即

$$Y = aX - a\mu_X + \mu_Y = aY + b.$$

6) 得证.

下一定义给出了协方差的一个无量纲描述形式.

定义 14　设 X 和 Y 是两个均值有限且方差非零的随机变量, 定义它们的相关系数如下

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}.$$

从定义中很快就能得到

$$-1 \leqslant \rho_{X,Y} \leqslant 1.$$

而且, 我们将看到 $\rho_{X,Y}$ 达到边界值 ± 1 当且仅当 X 与 Y 之间存在线性关系, 也就是说, 存在常数 $a \neq 0$ 和 b, 使得

$$Y = aX + b.$$

事实上, $\rho_{X,Y} = 1$ 暗示着斜率 $a \geqslant 0$, $\rho_{X,Y} = -1$ 则暗示着 $a \leqslant 0$. 因此, 如果 $\rho_{X,Y} = 1$, 则 Y 与 X 沿着相同的方向移动 (要么都增, 要么都减), 而当 $\rho_{X,Y} = -1$ 时, 则 X 增加时 Y 减少, 反之一样.

同样地我们容易看到, 当 X 与 Y 独立时, 有 $\rho_{X,Y} = 0$. 但是, 反过来不一定成立. 由 $\rho_{X,Y} = 0$ 不能推出 X 与 Y 独立.

称两个随机变量不相关, 如果 $\rho_{X,Y} = 0$. 称它们完全正相关, 如果 $\rho_{X,Y} = 1$; 称它们完全负相关, 如果 $\rho_{X,Y} = -1$. 研究投资组合的风险管理时会经常用到这些术语.

最佳线性估计

我们来更进一步地了解相关系数的含义. 经常说相关性度量了 X 与 Y 之间的线性关系. 确实, 我们已经看到随机变量完全相关等价于它们之间具有线性关系.

为了看得更深入一些, 假设用某线性函数 $\beta X+\alpha$ 来逼近 Y. 称这样的逼近为 Y 的最佳线性逼近. 称逼近误差

$$\varepsilon = Y - \beta X - \alpha$$

为剩余随机变量. 一般认为使得均值平方误差最小的线性估计为最合适的逼近, 均值平方误差的定义如下

$$\text{MSE} = \mathcal{E}(\varepsilon^2) = \mathcal{E}[(Y-\beta X-\alpha)^2].$$

当 $\rho_{X,Y}=\pm 1$ 时, 我们已经知道存在精确的逼近使得 $\text{MSE}=0$.

通常 MSE 能够写成如下形式

$$\text{MSE} = \mathcal{E}(Y^2) - 2\beta\mathcal{E}(XY) - 2\alpha\mathcal{E}(Y) + \beta^2\mathcal{E}(X^2) + 2\alpha\beta\mathcal{E}(X) + \alpha^2.$$

上式的最小值（一定存在）可通过使其导数为 0 而求得. 我们将求导过程留给读者, 从而得到方程

$$\begin{aligned}\beta\mathcal{E}(X^2) + \alpha\mathcal{E}(X) &= \mathcal{E}(XY),\\ \beta\mathcal{E}(X) + \alpha &= \mathcal{E}(Y).\end{aligned}$$

解之得

$$\begin{aligned}\beta &= \frac{\sigma_{X,Y}}{\sigma_X^2},\\ \alpha &= \mathcal{E}(Y) - \beta\mathcal{E}(X).\end{aligned}$$

下面我们给出总结, 首先从定义开始.

定义 15 设 X 和 Y 是随机变量, 令

$$Y = \beta X + \alpha + \varepsilon,$$

其中 β 和 α 是常数, ε 是如下定义的随机变量

$$\varepsilon = Y - \beta X - \alpha.$$

因此, 用线性函数 $\beta X+\alpha$ 来逼近 Y 时的随机误差为 ε. Y 相对于 X 的最佳线性估计, 记为BLP, 是使得均值平方误差 $\mathcal{E}(\varepsilon^2)$ 最小的线性函数 $\beta X+\alpha$. 称系数 β 为 Y 相对于 X 的贝塔值, 直线 $y=\beta x+\alpha$ 为回归线.

定理 9　Y 相对于 X 的最佳线性估计为

$$\mathrm{BLP} = \frac{\sigma_{X,Y}}{\sigma_X^2}X + \mu_Y - \frac{\sigma_{X,Y}}{\sigma_X^2}\mu_X.$$

此外, 最小均值平方误差为

$$\mathcal{E}(\varepsilon^2) = \sigma_Y^2(1 - \rho_{X,Y}^2).$$

下面叙述相关系数的一些性质：

- $\rho_{X,Y} = \pm 1$ 当且仅当 X 与 Y 之间存在线性关系.
- $\rho_{X,Y}$ 越接近于 ± 1, 最佳线性估计的均值平方误差越小.
- 如果 $\rho_{X,Y}$ 为正, 则 BLP 的斜率为正. 因此当 X 增加时, Y 的 BLP 也增加; 减少时也一样.
- 如果 $\rho_{X,Y} = -1$, 则 BLP 的斜率为负. 因此当 X 增加时, Y 的 BLP 减少, 反之一样.

值得一提的是, 强相关并不意味着因果关系. 这是由于随机变量 Y 的值与另一随机变量 X 的取值存在近似线性关系, 并不意味着 X 的变动将引起 Y 的变动. 这只是说, 两个随机变量有类似的表现. 比如, 在 20 世纪 90 年代早期, 个人电脑的销售量有显著增长, 再看汽车销售量, 仅仅因为两者之间存在正的相关性, 并不能说明个人电脑的购买将引起汽车的购买.

和的方差

我们需要得到随机变量线性组合的方差公式, 这时正好需要用上协方差. 定理 6 指出, 如果随机变量 X 和 Y 相互独立, 则有

$$\mathrm{var}(aX + bY) = a^2\mathrm{var}(X) + b^2\mathrm{var}(Y).$$

然而, 当随机变量不独立时, 上式并不成立. 这时我们有如下公式.

定理 10　设 X 和 Y 是 Ω 上的随机变量, $a, b \in R$, 则

$$\mathrm{var}(aX + bY) = a^2\mathrm{var}(X) + b^2\mathrm{var}(Y) + 2ab\mathrm{cov}(X, Y).$$

更一般地, 如果 $X_1, \cdots, X_n$ 是 Ω 上的随机变量, $a_1, \cdots, a_n$ 为常数, 则有

$$\mathrm{var}\Big(\sum_{i=1}^n a_i X_i\Big) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \mathrm{cov}(X_i, X_j).$$

练 习 1

1. 掷一对均匀的骰子, 求所得点数和为偶数的概率.

2. 掷 3 粒均匀的骰子, 求恰好一粒骰子的点数为 6 的概率.

3. 一个篮子里有 5 个红球, 3 个黑球, 4 个白球. 现随机地从篮子里摸出一个球.

a) 摸中红球的概率是多少?

b) 摸中白球或者红球的概率是多少?

c) 摸中的球不是红色的概率是多少?

4. 一个选择正确或者错误的测试包含 10 个问题. 一个学生随机地猜每个问题的答案.

a) 他猜中 10 个问题的概率是多少?

b) 他至少猜中 9 个问题的概率是多少?

c) 他至少猜中 8 个问题的概率是多少?

5. 一个骰子有六个面, 其中有两面的点数仅为 1, 其他四面的点数相应为 2, 3, 4, 5. 假设每面等可能性地出现.

a) 出现 1 点的概率是多少?

b) 出现 2 点的概率是多少?

c) 出现偶数点的概率是多少?

d) 出现的点数小于 3 的概率是多少?

6. 抛 4 枚均匀的硬币, 请问刚好出现 2 个正面的概率是多少?

7. 抛 4 枚均匀的硬币, 请问至少出现 2 个正面的概率是多少?

8. 掷一均匀的骰子, 同时从一副扑克里面随机抽取一张牌. 请问骰子出现的点数与所取牌上面的数字相配的概率是多少 (A 为一点)?

9. 通过对某一城市过去几十年天气情况的研究, 在三月份天气是晴或者有烟的概率如下:

P(晴天, 无烟) $= 0.07$, P(晴天, 轻烟) $= 0.09$,

P(晴天, 重烟) $= 0.12$, P(薄雾, 无烟) $= 0.09$,

P(薄雾, 轻烟) $= 0.07$, P(薄雾, 重烟) $= 0.11$,

P(阴天, 无烟) $= 0.16$, P(阴天, 轻雾) $= 0.12$,

P(阴天, 重烟) $= 0.17$.

问：晴天的概率是多少? 雾天的概率是多少? 没有烟或者轻烟的概率是多少?

10. a) 考虑一只股票, 它现在的价格为 50, 在固定时刻 T 的价格可能为 48, 49,

50, 51 中的一个. 假设股票可能价格的各自概率为

$$\mathbb{P}(48) = 0.2,$$
$$\mathbb{P}(49) = 0.4,$$
$$\mathbb{P}(50) = 0.3,$$
$$\mathbb{P}(51) = 0.1.$$

如果我们现在购买一份股票, 那么 T 时的期望收入是多少? 期望利润又是多少?

b) 考虑 a) 中这只股票的一衍生产品, 其收益 D 为股价的函数, 定义如下

$$D(48) = 2,$$
$$D(49) = -1,$$
$$D(50) = 0,$$
$$D(51) = 3.$$

因此, D 是 Ω 上的随机变量. 试问该衍生品的期望收益是多少?

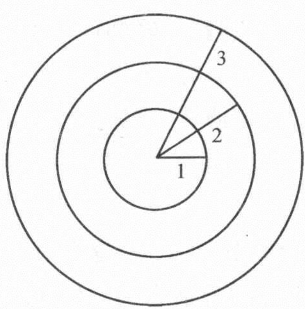
11. 假设你掷一均匀的骰子一次, 如果朝上的点数为偶数, 你赢得该数量的美元; 如果朝上的点数为奇数, 你失去该数量的美元. 玩这场赌博的期望收入是多少? 你愿意玩吗?

12. 花费 1 美元，你能掷一均匀的骰子一次。如果结果为偶数，你赢得 2 美元. 你愿意玩吗？为什么？

13. 假设你从一副扑克里面抽取一张牌, 如果该张牌为数字 (A 是 1), 你赢得相应数目的美元, 否则你失去 10 美元. 你的期望值是多少? 你愿意玩这样的赌博吗?

14. 一个美式的轮盘有 18 个红色数字, 18 个黑色数字和 2 个绿色数字. 如果你赌红色, 那么当真正出现红色数字时你将赢得和你本金一样多的钱 (同时拿回本金), 否则你将失去本金. 在该赌博中你的期望赢利是多少? 公平吗?

15. 考虑下面给出的靶图. 每支镖需花费 1.50 美元. 如果击中靶心你将得到 3 美元, 击中中环得到 2 美元, 击中外环得到 1 美元. 你的期望赢利是多少? 愿意玩这样的游戏吗?



16. 证明：$\text{var}(X) = \mathcal{E}(X)^2 - \mu^2$, 其中 $\mu = \mathcal{E}(X)$.

17. 证明：对任意的实数 a, 有

$$\text{var}(aX) = a^2 \text{var}(X).$$

18. 证明：如果随机变量 X 和 Y 相互独立, 则有

$$\text{var}(X+Y) = \text{var}(X) + \text{var}(Y).$$

19. 设随机变量 X 服从二项分布 $b(k;n,p)$, 证明 $\text{var}(X) = np(1-p)$. 提示：利用 $\text{var}(X) = \mathcal{E}(X^2) - \mathcal{E}(X)^2$.

20. 证明全概率定理.

21. 证明：如果随机变量 X, Y 和 Z 独立, 则 XY 与 Z 也独立.

22. 设 X 和 Y 是 $(\Omega, \mathbb{P})$ 上的独立随机变量, f 和 g 是从 $\mathbb{R}$ 到 $\mathbb{R}$ 的函数. 试证明 $f(X)$ 与 $g(Y)$ 独立.

23. 证明：$\rho_{X,Y} = +1$ 暗示着斜率 $a \geqslant 0$, $\rho_{X,Y} = -1$ 暗示着斜率 $a \leqslant 0$, 其中 $Y = aX + b$.

24. 证明：对任意的随机变量 X 和 Y, 有

$$\text{var}(aX+bY) = a^2\text{var}(X) + b^2\text{var}(Y) + 2ab\text{cov}(X,Y).$$

第 2 章　投资组合管理和资本资产定价模型

在这一章里, 我们来探讨资产组合的风险管理问题. 最主要的问题是如何去平衡一个投资组合, 也就是说, 在达到给定的期望收益下, 如何选择各资产的投资比例 (在价值上) 以使得风险最小化. 首先我们将认识到, 只知道投资组合中各资产的风险并不足以去了解投资组合的整个风险. 我们有必要考虑资产间的相互影响, 这是用单个风险之间的协方差 (或相关性) 来度量的.

2.1　投资组合、收益和风险

在我们的模型中, 假定只有两个时期: 初始时期 $t=0$ 和末期 $t=T$. 每种资产 a_i 的初始价值为 $\mathcal{V}_{i,0}$, 最终价值为 $\mathcal{V}_{i,T}$.

投资组合

一个投资组合是由资产 $a_i,\cdots,a_n$ 按一定比例形成的组合. 正式地, 我们定义一种投资组合为有序的 n 维实数

$$\Theta=(\theta_1,\cdots,\theta_n),$$

其中 θ_i 为资产 a_i 的投资比例. 如果 θ_i 为负, 表示投资组合在该资产上空头: 股票的卖空、卖权或买权的卖空等. θ_i 为正意味着多头: 持有某股票、买入卖权或买权等.

资产的权重

习惯性地, 用价值百分比来度量投资组合中某一资产所占数量. 资产 a_i 的权重 w_i 是 $t=0$ 时该资产在投资组合中所占的价值比例, 即

$$w_i=\frac{\theta_i\mathcal{V}_{i,0}}{\sum_{j=1}^{n}\theta_j\mathcal{V}_{j,0}}.$$

注意所有权重之和总为 1,

$$w_1+\cdots+w_n=1.$$

资产的收益

由如下方程来定义资产 a_i 的收益 (return)R_i,

$$\mathcal{V}_{i,T} = \mathcal{V}_{i,0}(1 + R_i),$$

等价于

$$R_i = \frac{\mathcal{V}_{i,T} - \mathcal{V}_{i,0}}{\mathcal{V}_{i,0}}.$$

因为资产在未来时刻 T 的价值是随机变量, 所以收益 R_i 也是随机变量. 因此我们可以考虑收益的期望和方差. 记资产 a_i 的期望收益为

$$\mu_i = \mathcal{E}(R_i).$$

方差为

$$\sigma_i^2 = \mathrm{var}(R_i).$$

称该方差为资产 a_i 的风险. 适当时我们也用标准差来度量风险.

投资组合的收益

投资组合本身的收益定义为各组成资产收益的加权和, 即

$$R = \sum_{i=1}^{n} w_i R_i.$$

例如, 假设一投资组合只含有两种资产, 其权重为 0.4 和 0.6, 相应的收益为 10% 和 8%. 那么该组合的收益为

$$(0.4)(0.10) + (0.6)(0.8) = 0.088 = 8.8\%.$$

由于期望算子是线性的, 故整个组合的期望收益为

$$\mu = \sum_{i=1}^{n} w_i \mu_i.$$

由于单个收益之间一般不独立, 故组合收益的方差由下式给出

$$\begin{aligned}\sigma^2 = \mathrm{var}\left(\sum_{i=1}^{n} w_i R_i\right) &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \mathrm{cov}(R_i, R_j)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \rho_{i,j}\sigma_i\sigma_j,\end{aligned}$$

其中 $\mathrm{cov}(R_i, R_j)$ 为 R_i 与 R_j 的协方差, $\rho_{i,j}$ 为相关系数.

定义 1　一个投资组合的期望收益 μ 为投资组合收益的期望值, 即

$$\mu = \mathcal{E}\Big(\sum_{i=1}^{n} w_i R_i\Big) = \sum_{i=1}^{n} w_i \mu_i.$$

定义组合的风险为组合收益的方差, 即

$$\begin{aligned}\sigma^2 == \operatorname{var}\Big(\sum_{i=1}^{n} w_i R_i\Big) &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \operatorname{cov}(R_i, R_j) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \rho_{i,j} \sigma_i \sigma_j.\end{aligned}$$

称资产是有风险的, 如果 σ_i^2 为正; 称资产是无风险的, 如果 σ_i^2 为 0.

除非特别提示, 我们假定投资组合中的所有资产都是有风险的, 即 $\sigma_i^2 > 0$.

对风险的进一步认识

我们来进一步认识风险这个概念. 一般来讲, 资产的风险有两种形式. 资产的系统风险是这样的风险, 它牵涉到整个市场的宏观面而不仅仅是某一特殊资产. 例如, 利率的改变会影响到整个市场. 另外一个例子就是一个国家货币供应量的改变. 诸如全球范围内的战争或恐怖活动可看成系统风险的一部分.

另一方面, 非系统风险或者独特的风险 (unique risk) 指的是某一资产或某一组资产所特有的风险. 例如, 一个投资者决定投资于一个生产大头帮的公司, 这就存在许多非系统风险. 比如, 顾客可能对大头帮失去兴趣, 或者生产大头帮的工厂可能被烧掉.

这两类风险的关键区别在于非系统风险能够被分散化而系统风险不能. 例如, 一个投资者可通过投资于所有的大头帮生产公司来减少或者消除大头帮工厂被烧掉的风险. 通过这种方法, 如果一个大头帮生产工厂被烧掉, 另外一个大头帮公司能弥补. 更一般地, 投资者可通过投资所有的玩具和游戏公司来减少大头帮受到冷漠的风险. 毕竟, 当你听见一个小孩说他厌倦买大头帮的时候, 难道这时他会把零用钱存银行吗?

风险如何相互影响之初步

为了看清单个资产在风险上的影响, 我们来考虑只包含具有期望收益 μ_1 和风险 σ_1^2 的单一资产 a_1 的投资组合. 很明显, 组合的全部风险也为 σ_1^2. 现在让我们增加另一资产 a_2 到该投资组合中. 假设第二种资产的期望收益为 μ_2, 风险为 σ_2^2.

假如资产 a_1 的权重为 t, 那么资产 a_2 的权重为 $1-t$. 因此，投资组合的期望收益为

$$\mu = t\mu_1 + (1-t)\mu_2.$$

其风险为

$$\sigma^2 = t^2\sigma_1^2 + (1-t)^2\sigma_1^2 + 2t(1-t)\rho_{1,2}\sigma_1\sigma_2.$$

如何比较组合风险与组合中单个资产的风险? 我们可以假设 $0 < \sigma_1 \leqslant \sigma_2$(必要的话可以调换数字顺序).

首先假设单个资产间不相关, 即 $\rho_{1,2} = 0$. 投资组合的风险为

$$\sigma^2 = t^2\sigma_1^2 + (1-t)^2\sigma_1^2 = (\sigma_1^2 + \sigma_2^2)t^2 - 2\sigma_2^2 t + \sigma_2^2.$$

关于 t 的二次图形在图 1 的左边. 简单求导可知最小风险发生在

$$t_m = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}.$$

这时风险为

$$\sigma_m^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}.$$

注意到

$$0 < \sigma_m^2 < \min\{\sigma_1^2, \sigma_2^2\}.$$

故最小风险为正, 但比任一单个资产的风险都小.

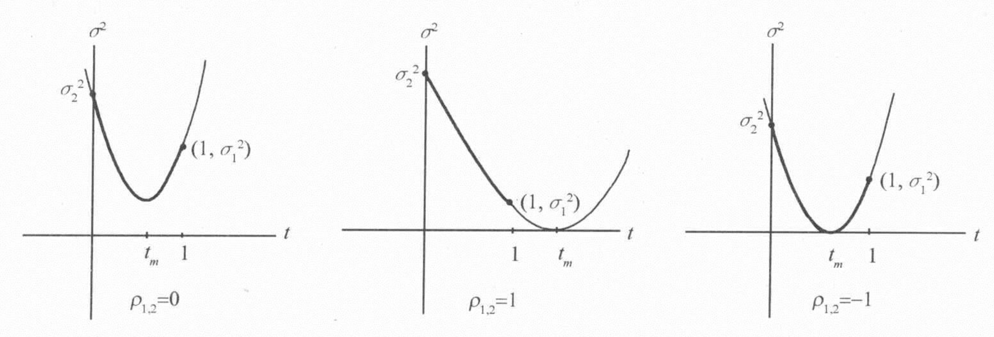


图 1　几种风险可能性 (黑体部分暗示不可卖空)

现在我们假设两单个资产完全正相关, 即 $\rho_{1,2} = 1$. 那么组合风险为

$$\sigma^2 = t^2\sigma_1^2 + (1-t)^2\sigma_1^2 + 2t(1-t)\sigma_1\sigma_2 = [(\sigma_1 - \sigma_2)t + \sigma_2]^2.$$

图 1 的中间显示了该二次图形. 最小风险确实为 0, 发生在

$$t_m = \frac{\sigma_2}{\sigma_2 - \sigma_1} > 0.$$

注意到

$$1 - t_m = -\frac{\sigma_1}{\sigma_2 - \sigma_1} < 0.$$

因此, 最小风险组合必须在具有较大风险的资产 a_2 上进行卖空 (持有空头).

最后, 假设两单个资产完全负相关, 即 $\rho_{1,2} = -1$. 那么组合风险为

$$\sigma^2 = t^2\sigma_1^2 + (1-t)^2\sigma_1^2 - 2t(1-t)\sigma_1\sigma_2 = [(\sigma_1 + \sigma_2)t - \sigma_2]^2.$$

图 1 的右边部分显示了该二次图形. 最小风险也为 0, 发生在

$$t_m = \frac{\sigma_2}{\sigma_2 + \sigma_1} > 0.$$

在该情形下

$$1 - t_m = \frac{\sigma_1}{\sigma_2 + \sigma_1} > 0.$$

所以最小风险组合不需要在任一资产上卖空.

因此, 单个资产之间完全负相关的情形看起来是最有希望的, 这种情形下组合风险可以降为 0 而不需卖空资产. 确实卖空有其自身缺点, 在很多情形下卖空需要额外的费用. 当然，在投资组合中要选择到完全负相关的资产一般非常困难 (或者不可能).

2.2 两种资产的投资组合

现在我们来认真地进行投资组合分析, 先从只包含两种资产 a_1 和 a_2 的投资组合开始, 组合里各资产的相应权重为 w_1 和 w_2. 通常习惯画出风险 – 期望收益曲线图, 其横坐标为风险, 纵坐标为期望收益. 在作图时通常用标准差作为风险度量.

这样一个投资组合的期望收益为

$$\mu = w_1\mu_1 + w_2\mu_2.$$

风险为

$$\sigma^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho_{1,2}\sigma_1\sigma_2.$$

像前面一样, 假定资产都是有风险的, 事实上, 可假定

$$0 < \sigma_1 \leqslant \sigma_2.$$

为了可读性, 记

$$\rho = \rho_{1,2}.$$

- $\rho = \pm 1$ 情形

首先考虑 $\rho = \pm 1$ 的情形. 这时简化 σ^2 的表达式可得到

$$\sigma = |w_1\sigma_1 \pm w_2\sigma_2|,$$

其中当 $\rho = 1$ 时取 "$+$", $\rho = -1$ 时取 "$-$". 因为 $w_1 + w_2 = 1$, 为了方便, 令

$$w_2 = s, \quad w_1 = 1 - s.$$

得到如下参数方程

$$\begin{aligned} \mu &= (1-s)\mu_1 + s\mu_2, \\ \sigma &= |(1-s)\sigma_1 \pm s\sigma_2|, \end{aligned}$$

这里 s 可取所有的实数. 当 s 取值于 $[0, 1]$ 上时, 两个权重都非负, 因此投资组合中无空头. 当 s 不在这个范围内时, 有一权重肯定为负, 意味着相应的资产被卖空而持有另一资产的多头.

为了帮助在平面上画出点 (σ, μ), 我们同时忽略绝对值符号而考虑下面的参数方程

$$\begin{aligned} \mu &= (1-s)\mu_1 + s\mu_2, \\ \sigma' &= (1-s)\sigma_1 \pm s\sigma_2. \end{aligned}$$

这些方程在 (σ', μ) 平面上是一条直线. 当 $\rho = 1$ 取 "$+$" 时, 该直线通过点 (σ_1, μ_1) 和 (σ_2, μ_2). $\rho = -1$ 时该直线通过点 $(-\sigma_1, \mu_1)$ 和 (σ_2, μ_2). 图 2 显示出了这些直线.

现在, 绝对值符号的作用就是简单地将直线位于左半平面里面的部分沿 μ轴 (因为 $\sigma = |\sigma'|$) 翻折过来. 结果见图 3. 黑体部分对应着两个权重都非负, 即没有卖空.

从参数方程 (或以前的讨论) 我们能推导出下面的定理, 该定理再次指出了我们能把组合风险降低到 0 的这些情形.

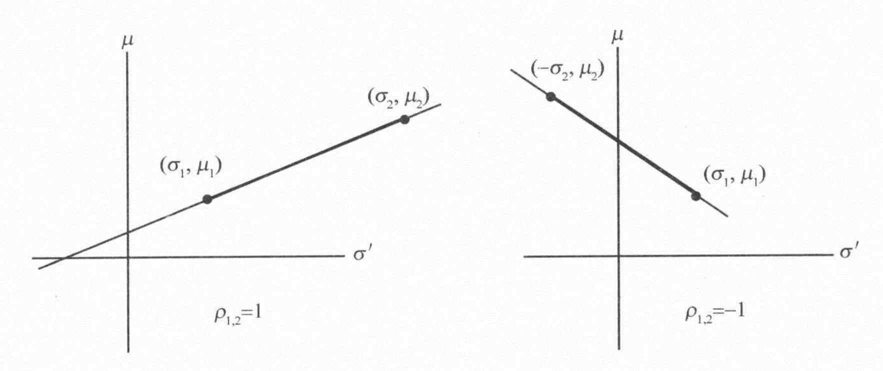


图 2　取绝对值前的图形

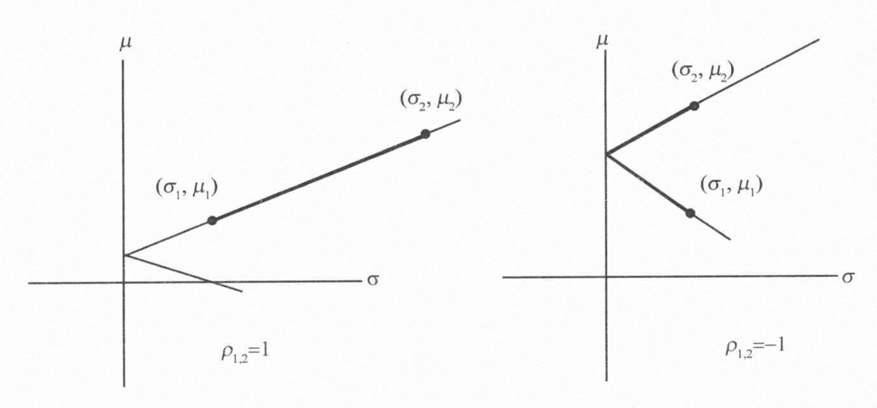


图 3　风险 – 收益线

定理 1　当 $\rho=\rho_{1,2}=\pm 1$ 时, 投资组合的期望收益和风险由如下的参数方程给出

$$\mu=(1-s)\mu_1+s\mu_2,$$
$$\sigma=|(1-s)\sigma_1\pm s\sigma_2|,$$

其中 s 为资产 a_2 所占的权重, 可取任意实数. 当 s 取值于 $[0,1]$ 上时, 两个权重都非负, 因此投资组合中无空头. 当 s 不在这个范围内时, 有一权重肯定为负, 意味着相应的资产被卖空而持有另一资产的多头. 图 3 给出了 (σ,μ) 的图形.

而且, 我们有下面这些情形:

1) 当 $\rho=1$ 和 $\sigma_1=\sigma_2$ 时, 所有的权重可能组合形成的投资组合都有相同的风险 (因此也是最小风险)$\sigma_{\min}=\sigma_1=\sigma_2$.

2) 当 $\rho=1$ 和 $\sigma_1<\sigma_2$ 时, 最小风险组合的各权重为

$$w_1=\frac{-\sigma_2}{\sigma_1-\sigma_2},\quad w_2=\frac{\sigma_1}{\sigma_1-\sigma_2}.$$

同时有

$$\mu_{\min} = \frac{\sigma_1\mu_2 - \sigma_2\mu_1}{\sigma_1 - \sigma_2},$$
$$\sigma_{\min} = 0.$$

3) 当 $\rho = -1$ 时, 最小风险组合的各权重为

$$w_1 = \frac{\sigma_2}{\sigma_1 + \sigma_2}, \quad w_2 = \frac{\sigma_1}{\sigma_1 + \sigma_2}.$$

同时有

$$\mu_{\min} = \frac{\sigma_1\mu_2 + \sigma_2\mu_1}{\sigma_1 + \sigma_2},$$
$$\sigma_{\min} = 0.$$

我们要强调的一点是, 一般不太可能找到具有完全相关性的资产, 所以上面的定理仅仅是理论上的结论. 它指出了组合资产的风险对各单个资产之间的相关系数的依赖.

- $-1 < \rho < 1$情形

当 $-1 < \rho < 1$ 时, 期望收益和风险的参数方程为

$$\mu = w_1\mu_1 + w_2\mu_2,$$
$$\sigma^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho\sigma_1\sigma_2.$$

现将上面的式子参数化, 令

$$w_2 = s, \quad w_1 = 1 - s.$$

得到

$$\mu = (\mu_2 - \mu_1)s + \mu_1,$$
$$\sigma^2 = (\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)s^2 - 2\sigma_1(\sigma_1 - \rho\sigma_2)s + \sigma_1^2.$$

接下来观察到, 因为 $\rho < 1$, 所以 σ^2 中 s^2 的系数满足

$$\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2 = (\sigma_1 - \sigma_2)^2 + 2(1-\rho)\sigma_1\sigma_2 > 0.$$

因此 σ^2 的表达式确实是 s 的二次函数 (非线性). 由点 (σ^2, μ) 形成的图形是开口向右的抛物线, 图形经过点 (σ_1^2, μ_1) 和 (σ_2^2, μ_2). 图 4 给出了点 (σ_1^2, μ_1) 和 (σ_2^2, μ_2) 两种可能位置. 在右边的这个图里, 最小风险组合要求一些资产空头.

再次假设 $0 < \sigma_1 \leqslant \sigma_2$. 将 σ^2 对 s 求导得到

$$\frac{\mathrm{d}}{\mathrm{d}s}(\sigma^2) = 2(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)s - 2\sigma_1(\sigma_1 - \rho\sigma_2).$$

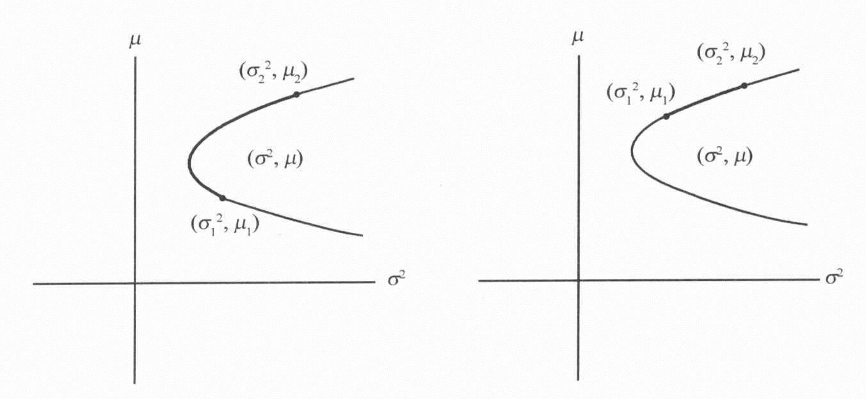


图 4　风险－收益图

因此最小风险发生在

$$s_{\min} = \frac{\sigma_1(\sigma_1 - \rho\sigma_2)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}.$$

这时最小风险为

$$\sigma_{\min}^2 = \frac{\sigma_1^2\sigma_2^2(1-\rho^2)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}.$$

最小风险组合里的各资产都不需要卖空当且仅当 $s_{\min} \in [0,1]$.

我们能够结合 $\rho = \pm 1$ 情形, 如果排除 $\rho = 1, \sigma_1 = \sigma_2$ 这种退化的情形, 因为该情形时分母为 0. 既然退化的情形被排除了, 那么稍微知道一点代数知识就能得到 $s_{\min} < 1$(退化情形时相应的 $s_{\min} = 1$). 而且

$$\begin{aligned} 0 < s_{\min} < 1 \quad &\Leftrightarrow \quad -1 \leqslant \rho < \frac{\sigma_1}{\sigma_2}, \\ s_{\min} = 0 \quad &\Leftrightarrow \quad \rho = \frac{\sigma_1}{\sigma_2}, \\ s_{\min} < 0 \quad &\Leftrightarrow \quad \frac{\sigma_1}{\sigma_2} < \rho \leqslant 1. \end{aligned}$$

这是最后的结果.

定理 2　假设 $0 < \sigma_1 \leqslant \sigma_2$, 令 $\rho = \rho_{1,2}$ 为相关系数. 更进一步假定, 如果 $\rho = 1$, 那么 $\sigma_1 \neq \sigma_2$. 如果用 $s_{\min}$ 来表示风险最下化时资产 a_2 的权重, 那么

$$s_{\min} = \frac{\sigma_1(\sigma_1 - \rho\sigma_2)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2},$$

且

$$\begin{aligned} \mu_{\min} &= (\mu_2 - \mu_1)s_{\min} + \mu_1, \\ \sigma_{\min}^2 &= \frac{\sigma_1^2\sigma_2^2(1-\rho^2)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}. \end{aligned}$$

进一步有

1) 当 $\rho=1$ 和 $\sigma_1=\sigma_2$ 时, 所有的可能权重组合形成的投资组合都有相同的风险(因此也是最小风险)为

$$\sigma_{\min}=\sigma_1=\sigma_2.$$

2) 条件 $-1\leqslant\rho<\sigma_1/\sigma_2$ 与 $0<s_{\min}<1$ 等价, 因此最小风险组合可以不需要卖空而实现. 而且

$$\sigma_{\min}^2<\min\{\sigma_1^2,\sigma_2^2\},$$

但是

$$\sigma_{\min}^2\Leftrightarrow\rho=-1.$$

3) 条件 $\rho=\sigma_1/\sigma_2$(除了 $\rho=1,\sigma_1=\sigma_2$ 情形) 与 $s_{\min}=0$ 等价, 在该情形下,

$$\sigma_{\min}^2=\sigma_1^2$$

可仅仅持有资产 a_1 而实现.

4) 条件 $\sigma_1/\sigma_2<\rho\leqslant 1$ 等价于 $s_{\min}<0$, 因此需卖空资产 a_2 来使风险最小化. 而且

$$\sigma_{\min}^2<\min\{\sigma_1^2,\sigma_2^2\},$$

但是

$$\sigma_{\min}^2=0\Leftrightarrow\rho=1.$$

与前面的那个定理一样, 该结果仅仅是理论上的结论, 但是它指出了组合资产的风险对各单个资产之间的相关系数的依赖.

2.3 多资产的投资组合

现在我们将注意力转移到至少包含两种资产的投资组合上. 组合中的权重可以表示成如下矩阵 (或向量) 形式

$$W=(w_1\ \ w_2\ \ \cdots\ \ w_n).$$

也很方便定义元素全为 1 的矩阵 (或向量) 如下

$$O=(1\ \ 1\ \ \cdots\ \ 1).$$

上面的标记 “O” 代表 “1”, 因此条件

$$w_1 + \cdots + w_n = 1$$

可写成矩阵的乘积形式

$$OW^{\mathrm{T}} = 1,$$

这里 W^{T} 表示 W 的转置. 我们也将期望收益矩阵写成

$$M = (\mu_1 \;\; \mu_2 \;\; \cdots \;\; \mu_n).$$

将协方差矩阵写成

$$C = (c_{i,j}),$$

其中

$$c_{i,j} = \mathrm{cov}(R_i, R_j).$$

注意到 $c_{i,i} = \sigma_i^2$ 为 R_i 的方差. 这里 C 是对称矩阵 (即 $C^{\mathrm{T}} = C$) 且是半正定的, 意思是对任意矩阵 $A = (a_1, \cdots, a_n)$, 有 $ACA^{\mathrm{T}} \geqslant 0$. 我们也假定 C 是可逆的, 这就意味着 C 是正定的, 即对任意矩阵 $A = (a_1, \cdots, a_n)$, 有 $ACA^{\mathrm{T}} > 0$.

期望收益和方差能写成如下矩阵乘积形式

$$\begin{aligned} \mu &= MW^{\mathrm{T}} = \mu_1 w_1 + \cdots + \mu_n w_n, \\ \sigma^2 &= \mathrm{var}(R_1 w_1 + \cdots + R_n w_n) = \sum_{i,j=1}^{n} c_{i,j} w_i w_j = WCW^{\mathrm{T}}. \end{aligned}$$

马柯维茨弹头 (Markowitz bullet)

我们来考察投资组合中各资产权重 $W = (w_1 \;\; w_2 \;\; \cdots \;\; w_n)$ 与相应风险 - 期望收益 (σ, μ) 之间的关系。注意我们现在是用标准差 σ 在度量风险.

图 5 详细地描绘了 3 种资产的情形, 这为一般情况下多种资产情形提供了几何直观感受 (我们稍后将定义马柯维茨弹头和马柯维茨有效前沿).

图 5 的左边部分表示出了一个 n 维空间, 权重向量 $(w_1, \cdots, w_n)$ 落在该空间上 (当然图 5 显示的是 $n = 3$ 时的情形). 由于各权重之和肯定为 1, 因此权重向量必须落在如下给出的超平面上

$$w_1 + \cdots + w_n = 1.$$

$n=3$ 时的超平面是 3 维空间上的平常意义上的平面, 通过点 $(1,0,0),(0,1,0)$和 $(0,0,1)$. 由于透明的缘故, 该图只显示了超平面位于正值区域内的部分. 这部分对应着投资组合中资产没有卖空的情形. 下面我们把整个超平面看作权重超平面 (weight hyperplane).

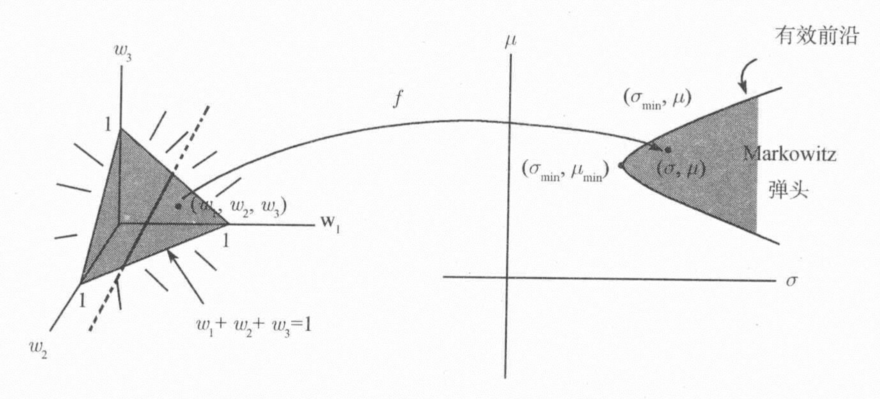


图 5 Markowitz 弹头

我们把权重超平面上的权重向量到相应投资组合的风险 – 期望收益有序对之间的对应关系用函数 f 来表示, 即

$$f(w_1,\cdots,w_n)=(\sigma,\mu),$$

其中

$$\mu=\mu_1 w_1+\cdots+\mu_n w_n=MW^{\mathrm{T}},$$

$$\sigma^2=\sum_{i,j=1}^{n} c_{i,j}w_i w_j=WCW^{\mathrm{T}}.$$

函数 f 也在图 5 中显示了出来. 我们的目标就是确定权重超平面上的一条直线在 f 下的象. 这能帮助我们理解 f 在一般情形下是如何作用的 (与作函数的边界图类似).

n 维空间上 (不管是权重超平面还是其他) 的一条线的方程可写成如下参数形式

$$l(t)=(a_1 t+b_1,\cdots,a_n t+b_n)=At+B,$$

其中

$$A=(a_1,\cdots,a_n),$$

$$B=(b_1,\cdots,b_n).$$

参数 t 的取值范围为 $(-\infty, \infty)$. $t = 0$ 对应着点 $l(0) = B$, $t = 1$ 对应着 $l(1) = A + B$. 任何这种形式的方程都是一条线的方程.

现在, 线上的任意一点 $W = (w_1, \cdots, w_n)$ 都对应着一特定的 t, 其期望收益为

$$\mu = MW = M(At + B) = (MA)t + MB.$$

可以看出期望收益是 t 的函数. 这一点非常重要. 求解 t 得到

$$t = \frac{\mu - MB}{MA} = \alpha\mu + \beta,$$

这里用 α 和 β 表示是为了方便, 当然必须假设分母 MA 不为 0.

现在让我们来看看风险 (以方差的形式)

$$\begin{aligned}\sigma^2 &= WCW^{\mathrm{T}} \\ &= (At + B)C(A^{\mathrm{T}}t + B^{\mathrm{T}}) \\ &= (ACA^{\mathrm{T}})t^2 + (BCA^{\mathrm{T}} + ACB^{\mathrm{T}})t + BCB^{\mathrm{T}} \\ &= \gamma t^2 + \delta t + \varepsilon,\end{aligned}$$

这里我们用 $\gamma, \delta, \varepsilon$ 来简化表达式, 该表达式刚好是 t 的二次函数. 将 t 用 μ 的表达式代入, 得

$$\sigma^2 = \gamma(\alpha\mu + \beta)^2 + \delta(\alpha\mu + \beta) + \varepsilon$$

为 μ 的二次函数.

因此, 当 t 取遍 $(-\infty, \infty)$ 时, $l(t)$ 绘出权重超平面上的一条线, 风险 – 期望收益点 (σ^2, μ) 在 (σ, μ) 平面上形成了一条抛物线. 将点 (σ^2, μ) 中的第一个坐标取成平方根得到一条曲线, 称该曲线为马柯维茨曲线, 虽然这一术语并不标准. 因此, 权重超平面上的直线经 f 映射到 (σ, μ) 平面上的马柯维茨曲线. 注意马柯维茨曲线不是抛物线.

图 6 给出了由 Excel 软件绘出的一条马柯维茨曲线. 为了以后参考需要, 我们注意到用来绘制该曲线的点为

$$\begin{aligned}&(\mu_1, \mu_2, \mu_3) = (0.1, 0.11, 0.07), \\ &(\sigma_1, \sigma_2, \sigma_3) = (0.23, 0.26, 0.21), \\ &\rho_{1,2} = \rho_{2,1} = -0.15, \\ &\rho_{1,3} = \rho_{3,1} = 0.25, \\ &\rho_{2,3} = \rho_{3,2} = 0.2.\end{aligned}$$



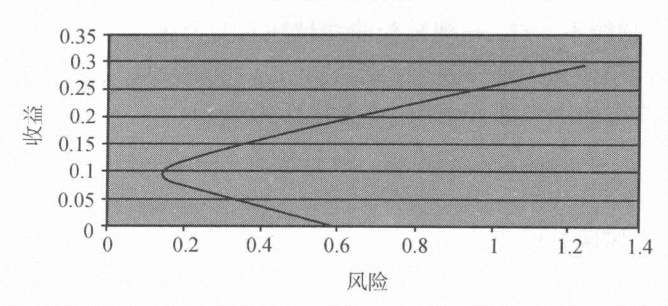


图 6 一种 Markowitz 弹头

马柯维茨曲线的形状

非常重要的是, 必须分清由 (σ^2, μ) 形成的抛物线和由 (σ, μ) 形成的马柯维茨曲线 (图 6) 之间的差别. 为了对这种差别有好的理解, 考虑如下函数 $(a > 0)$

$$y = ax^2 + bx + c$$

和

$$z = \sqrt{ax^2 + bx + c}\,.$$

第一个图像是抛物线. 通过求导可知抛物线上的切线斜率方程为

$$y' = 2ax + b$$

随 x 趋于 ∞ 时, 斜率逐步递增到 ∞(无界). 另一方面, 对函数 z 来说, 当 x 比较大时, 第一项占主要地位, 因此

$$z = \sqrt{ax^2 + bx + c} \approx \sqrt{ax^2} = \sqrt{a}|x|.$$

方程 $z = \sqrt{a}|x|$ 的图像是一对射线. 这说明了当 x 趋于 ∞ 时, z 的图像逐渐变平, 不像抛物线的情形. 特别, z 的导数为

$$z' = \frac{2ax + b}{2\sqrt{ax^2 + bx + c}}.$$

为了容易些, 我们先平方再求极限, 即有

$$\lim_{x\to\infty} (z')^2 = \lim_{x\to\infty} \frac{4a^2x^2 + 4abx + b^2}{4(ax^2 + bx + c)} = a.$$

故当 x 趋于 ∞ 时, z' 趋于 $\sqrt{a}$.

因此, 与抛物线不一样, 马柯维茨曲线越往右变得越平. 这个事实的一个暗示, 对资本资产定价模型来说非常重要, 就是当 μ_{rf} 特别大时, 过点 (o, μ_{rf}) 不存在直线与马柯维茨曲线的上半部分相切 (可看后面的图 9).

最小风险点

我们记最小风险处的点为 $(\sigma_{\min}, \mu_{\min})$. 下面将寻找与该点相对应的组合权重 (在权重超平面上). 对于任何特殊的情形, 将权重代入 σ 和 μ 的公式而容易得到精确点 (读者将看到, 一般的公式可能有点杂乱).

下面的定理给出了与最小风险相应的权重. 其证明主要是利用拉格朗日乘子, 有关拉格朗日乘子能在任一本多变量微积分的著作中找到, 因此我们在这里就不介绍了. 如果有必要的话, 读者可跳过部分需要用到拉格朗日乘子的地方.

定理 3　*具有最小风险的投资组合的权重为*

$$W = \frac{OC^{-1}}{OC^{-1}O^{\mathrm{T}}}.$$

注意, 分母是一个数, 刚好是分子中的各分量之和.

证明　我们来求解如下问题

$$\begin{aligned} \min\ \sigma^2 &= \sum_{i,j=1}^{n} c_{i,j} w_i w_j = WCW^{\mathrm{T}}, \\ \text{s.t.}\quad OW^{\mathrm{T}} &= w_1 + \cdots + w_n = 1. \end{aligned}$$

根据拉格朗日乘子方法, 我们必须将如下函数对每个 w_i 和 α 求导

$$g(w_1, \cdots, w_n) = \sum_{i,j=1}^{n} c_{i,j} w_i w_j + \alpha(1 - w_1 - \cdots - w_n).$$

令各导数为 0. 我们将这些过程作为练习, 从而得到

$$CW^{\mathrm{T}} = \frac{\alpha}{2} O^{\mathrm{T}},$$

故

$$W = \frac{\alpha}{2} OC^{-1}.$$

将其代入约束条件 (并利用 C, C^{-1} 都是对称的), 得到

$$\frac{\alpha}{2} = \frac{1}{OC^{-1}O^{\mathrm{T}}}.$$

得证.

马柯维茨有效前沿

点 $(\sigma_{\min}, \mu)$ 表示在给定期望收益 μ 下, 最小风险为 $\sigma_{\min}$, 称由所有这样的点组成的集合为马柯维茨有效前沿 (“前沿”是边界的另一说法). 下面的定理描述了这些点的集合. 虽然公式有点繁乱, 但非常重要. 也就是, 最小风险的权重是期望收益的线性函数. 意思是当期望收益 μ 取遍 $-\infty$ 到 ∞ 时, 最小风险的权重轨迹为权重超平面上的一条直线, 且相应点 $(\sigma_{\min}, \mu)$ 的轨迹正是马柯维茨曲线.

换句话说, 马柯维茨有效前沿是一条马柯维茨曲线. 称对应着马柯维茨曲线的权重线为最小风险权重线 (minimum-risk weight line).

定理 4　对一给定的期望收益 μ, 具有最小风险的投资组合中的资产权重为

$$W = \frac{\begin{vmatrix} \mu & MC^{-1}O^{\mathrm{T}} \\ 1 & OC^{-1}O^{\mathrm{T}} \end{vmatrix} MC^{-1} + \begin{vmatrix} MC^{-1}M^{\mathrm{T}} & \mu \\ OC^{-1}M^{\mathrm{T}} & 1 \end{vmatrix} OC^{-1}}{\begin{vmatrix} MC^{-1}M^{\mathrm{T}} & MC^{-1}O^{\mathrm{T}} \\ OC^{-1}M^{\mathrm{T}} & OC^{-1}O^{\mathrm{T}} \end{vmatrix}}.$$

特别地, 每个权重 w_i 都是 μ 的线性函数.

证明　在该情形下, 我们求解如下问题

$$\begin{aligned} &\min\ \sigma^2 = \sum_{i,j=1}^{n} c_{i,j} w_i w_j = WCW^{\mathrm{T}}, \\ &\text{s.t.}\quad MW^{\mathrm{T}} = w_1\mu_1 + \cdots + w_n\mu_n = \mu, \\ &OW^{\mathrm{T}} = w_1 + \cdots + w_n = 1. \end{aligned}$$

通过令如下函数的偏导数为 0 而求得

$$g = \sum_{i,j=1}^{n} c_{i,j} w_i w_j + \alpha(\mu - w_1\mu_1 - \cdots - w_n\mu_n) + \beta(1 - w_1 - \cdots - w_n).$$

解方程组可得

$$2CW^{\mathrm{T}} = \alpha M^{\mathrm{T}} + \beta O^{\mathrm{T}}.$$

因此

$$W = \frac{1}{2}(\alpha M + \beta O)C^{-1}.$$

将其代入约束条件, 得

$$\begin{aligned} (MC^{-1}M^{\mathrm{T}})\alpha + (MC^{-1}O^{\mathrm{T}})\beta &= 2\mu, \\ (OC^{-1}M^{\mathrm{T}})\alpha + (OC^{-1}O^{\mathrm{T}})\beta &= 2. \end{aligned}$$

利用克拉默法则(Cramer's rule)可求得 α 和 β. 再代入 W 中即可. 我们将证明过程中的一些必要细节留给读者作为练习.

称一有序对 (x, y) 为可达到点 (attainable point), 如果它具有某一投资组合的 (σ, μ) 形式. 既然马柯维茨有效前沿包含了所有风险最小的点, 因此所有的可达到点定会落入前沿上某点的右边 (对应于更大的风险). 换句话说, 可达到点在图 5 右边的阴影部分内. 根据其形状, 该区域 (包括边界) 被叫做马柯维茨弹头.

为了解释马柯维茨有效前沿的重要性, 我们先给出下面的定义.

定义 2 设 $P_1 = (\sigma_1, \mu_1), P_2 = (\sigma_2, \mu_2)$ 为可达到点. 称 (σ_1, μ_1) 占优 (σ_2, μ_2), 如果

$$\sigma_1 \leqslant \sigma_2, \quad 且\mu_1 \geqslant \mu_2.$$

简单明了地说, P_1 有更小或相同的风险和更大或相等的期望收益.

定理 5 任一可达到点都被马柯维茨有效前沿上的某一点占优. 因此, 对任何期望收益都追求风险最小的投资者来说, 他只需要在马柯维茨有效边界上寻找投资组合.

例 1 我们跳过一些必要的计算而得到图 6 中的马柯维茨弹头. 数据如下

$$\begin{aligned}
&(\mu_1, \mu_2, \mu_3) = (0.1, 0.11, 0.07),\\
&(\sigma_1, \sigma_2, \sigma_3) = (0.23, 0.26, 0.21),\\
&\rho_{1,2} = \rho_{2,1} = -0.15,\\
&\rho_{1,3} = \rho_{3,1} = 0.25,\\
&\rho_{2,3} = \rho_{3,2} = 0.2.
\end{aligned}$$

因为计算是乏味的, 所以最好用软件程序来处理, 比如Excel. 图 7 给出了一张Excel表格的一部分, 上面有需要的计算结果. 用户只需填灰色的单元, 剩余部分能够自动调整. 参照图 7, 由定理 3 给出的最小风险的点的资产权重为

$$W = \frac{OC^{-1}}{OC^{-1}O^{\mathrm{T}}}.$$

利用权重就得到期望收益和风险为

$$\begin{aligned}
\mu &= MW^{\mathrm{T}},\\
\sigma^2 &= WCW^{\mathrm{T}}.
\end{aligned}$$

因此, 具有最小风险的点为

$$(\sigma_{\min}, \mu_{\min}) = (0.146, 0.094) = (14.6\%, 9.4\%).$$

接下来, 我们计算达到一给定期望收益 μ 的投资组合的最小风险. 求最小风险的公式在定理 4 中已经给出. 在图 7 中计算了所需的全部矩阵乘积. 图 8 给出了不同期望收益下的最小风险的计算结果.

资本资产定价模型: 灰色单元为输入区

用户数据		收益 μ_i	风险 σ_i	相关性	
	i=1	0.1	0.23	−0.15	$=\rho_{1,2}=\rho_{2,1}$
	i=2	0.11	0.26	0.25	$=\rho_{1,3}=\rho_{3,1}$
	i=3	0.07	0.21	0.2	$=\rho_{2,3}=\rho_{3,2}$

$C=(c_{i,j})=(\rho_{i,j}\sigma_i\sigma_j)$		j=1	j=2	j=3
	i=1	0.0529	−0.00897	0.012075
	i=2	−0.00897	0.0676	0.01092
	i=3	0.012075	0.01092	0.0441
C的逆		21.10168374	3.888932099	−6.740815639
		3.888932099	16.12598046	−5.05792657
		−6.740815639	−5.05792657	25.77387544

最小风险点	$OC^{-1}=$	18.2498002	14.95698599	13.97513323
	$OC^{-1}O^{\mathrm{T}}=$	47.18191942		
	$W=$	0.386796477	0.31700673	0.296196793
	$\mu=$	0.094284164		
	$WC=$	0.02119456	0.02119456	0.02119456
	$WCW^{\mathrm{T}}=$	0.02119456		
	$\sigma=$	0.145583514		
最小风险线	$MC^{-1}=$	2.06609381	1.808696201	0.573717794
	$MC^{-1}O^{\mathrm{T}}=$	4.448507805		
	$OC^{-1}M^{\mathrm{T}}=$	4.448507805		
	$MC^{-1}M^{\mathrm{T}}=$	0.445726209		
	Denom Det=	1.24099637		

图 7 Excel 工作表格

收益 μ	Num Det 1	Num 1st Term		
0	−4.448507805	−9.19103444	−8.045999165	−2.55218809
0.005	−4.212598208	−8.703623082	−7.619310373	−2.41684255
0.01	−3.976688611	−8.216211723	−7.192621581	−2.28149702

Num Det 2	Num 2nd Term		
0. 445726209	8.134414252	6.666720658	6.229083152
0. 42348367	7.728492359	6.334039314	5.918240706
0. 401241131	7.322570466	6.001357969	5.607398259

Num of W			W		
−1.056620188	−1.379278507	3.676895066	−0.851428911	−1.111428317	2.962857228
−0.975130723	−1.28527106	3.501398153	−0.785764363	−1.035676727	2.821441091
−0.893641258	−1.191263612	3.32590124	−0.720099816	−0.959925138	2.680024954

收益 μ	WC			$\sigma_2=WCW^{\mathrm{T}}$	风险 σ
1.11022E-16	0.000705424	−0.035140836	0.108244202	0.35916802	0.599306
0.005	0.001791987	−0.032153304	0.103627858	0.324272245	0.569449
0.01	0.00287855	−0.029165771	0.099011513	0.291277439	0.539701

图 8 在一给定期望收益下计算最小风险

资本资产定价模型 (CAPM)

我们已经讨论了马柯维茨投资组合理论, 接下来看看资本资产定价模型. 将马柯维茨投资组合理论演变到资本资产定价模型最主要的一点就是在模型中引进了无风险资产 (riskfree asset)(回忆一下, 直到现在, 我们一直假定所有资产都是有风险的).

如我们已经说过的, 无风险资产就是风险为 0, 即方差为 0 的资产. 因此它的风险 – 期望收益点落在纵轴上, 如图 9 所示.

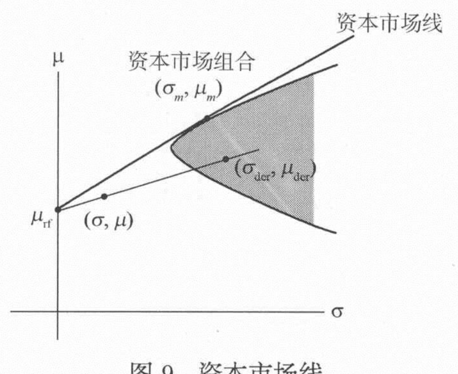


图 9　资本市场线

一般认为将无风险资产加入到马柯维茨投资组合理论中去是 William Sharp 的贡献, 他因此获得了诺贝尔奖. 但是 John Lintner 和 Mossin 也在同一时候独立地得到了相似的结果. 由于这些原因, 该理论有时被叫做 Sharp-Lintner-Mossin (SLM) 资本资产定价模型.

CAPM 里面最基本的思想就是一个投资者能够改进他的风险 / 期望收益平衡, 这可通过部分投资于风险资产的投资组合, 部分投资于无风险资产上来实现. 我们来看看为什么这是真的.

设想一个投资组合由权重为 w_{rf} 的无风险资产 a_{rf} 和权重分别为 $w_1, \cdots, w_n$ 的风险资产 $a_1, \cdots, a_n$ 组成. 注意现在风险资产的权重和至多为 1. 事实上, 我们有

$$w_{\mathrm{rf}} + \sum_{i=1}^{n} w_i = 1,$$

$$w_{\mathrm{risky}} = \sum_{i=1}^{n} w_i \leqslant 1.$$

整个投资组合的期望收益为

$$\mu = w_{\mathrm{rf}}\mu_{\mathrm{rf}} + \sum_{i=1}^{n} w_i\mu_i = w_{\mathrm{rf}}\mu_{\mathrm{rf}} + \mu_{\mathrm{risky}}.$$

因为无风险资产的方差为 0, 所以收益 $R_{\rm rf}$ 为一常数. 因此, 它与其他任何收益的协方差都为 0, 故

$$\begin{aligned}\sigma^2 &= \mathrm{var}\Big(w_{\rm rf}R_{\rm rf} + \sum_{i=1}^n w_iR_i\Big) = \mathrm{var}\Big(\sum_{i=1}^n w_iR_i\Big)\\ &= \sigma^2_{{\rm risk}y}.\end{aligned}$$

因此有

$$\sigma = \sigma_{{\rm risk}y}.$$

我们也想考虑这样的投资组合, 就是将无风险资产移走, 同时给每个风险资产的权重乘以同一因数而使得其和为 1. 就让我们称这样的投资组合为导出的风险组合 (derived riky portfolio) (一个非标准的术语). 例如, 如果原来投资组合的组成如下

$$\begin{aligned}\text{无风险资产的权重}w_{\rm rf} &= 0.20,\\ \text{风险资产}a_1\text{的权重}w_1 &= 0.30,\\ \text{风险资产}a_2\text{的权重}w_2 &= 0.50.\end{aligned}$$

因此风险资产的权重为 0.80, 故导出的风险组合包括

$$\begin{aligned}\text{风险资产}a_1\text{的权重}w_1 &= 0.30/0.80 = 0.375,\\ \text{风险资产}a_2\text{的权重}w_2 &= 0.50/0.80 = 0.625.\end{aligned}$$

总权重为 1. 我们记导出的风险组合的期望收益为 $\mu_{\rm der}$, 风险为 $\sigma^2_{\rm der}$. 那么有

$$\begin{aligned}\mu &= w_{\rm rf}\mu_{\rm rf} + \sum_{i=1}^n w_i\mu_i\\ &= w_{\rm rf}\mu_{\rm rf} + w_{{\rm risk}y}\sum_{i=1}^n \frac{w_i}{w_{{\rm risk}y}}\mu_i\\ &= w_{\rm rf}\mu_{\rm rf} + w_{{\rm risk}y}\mu_{\rm der}\end{aligned}$$

和

$$\begin{aligned}\sigma^2 &= \mathrm{var}\Big(\sum_{i=1}^n w_iR_i\Big)\\ &= w^2_{{\rm risk}y}\mathrm{var}\Big(\sum_{i=1}^n \frac{w_i}{w_{{\rm risk}y}}R_i\Big)\\ &= w^2_{{\rm risk}y}\sigma^2_{\rm der}.\end{aligned}$$

因此

$$\mu = w_{\text{rf}}\mu_{\text{rf}} + w_{\text{risky}}\mu_{\text{der}},$$
$$\sigma = w_{\text{risky}}\sigma_{\text{der}},$$

或者由 $w_{\text{rf}} + w_{\text{risky}} = 1$, 有

$$\mu = \mu_{\text{rf}} + w_{\text{risky}}(\mu_{\text{der}} - w_{\text{rf}}),$$
$$\sigma = w_{\text{risky}}\sigma_{\text{der}}.$$

当 w_{risky} 取遍所有的实数时, 上述方程组的轨迹为一条直线. 通过第二个方程解出 w_{risky}, 再代入第一个方程中去, 得到方程

$$\mu = \mu_{\text{rf}} + \frac{\mu_{\text{der}} - \mu_{\text{rf}}}{\sigma_{\text{der}}}\sigma.$$

图 9 给出了该直线.

显然, 如果 $w_{\text{risky}} = 0$, 则

$$(\sigma, \mu) = (0, \mu_{\text{rf}}).$$

如果 $w_{\text{risky}} = 1$, 则

$$(\sigma, \mu) = (\sigma_{\text{der}}, \mu_{\text{der}}).$$

而且, $w_{\text{risky}} = 1$ 对应的点 $(\sigma_{\text{der}}, \mu_{\text{der}})$ 为纯风险组合的风险 / 期望收益点, 该点一定落在马柯维茨弹头里.

那么我们站在哪里? 一个既在无风险资产又在一些风险资产上投资的投资者的风险 / 期望收益点应落在连接点 $(0, \mu_{\text{rf}})$ 和 $(\sigma_{\text{der}}, \mu_{\text{der}})$ 的直线上的某一位置上. 但是, 从几何图形上看很明显的是, 在所有连接点 $(0, \mu_{\text{rf}})$ 和位于马柯维茨弹头里的点 $(\sigma_{\text{der}}, \mu_{\text{der}})$ 的直线中, 在给定风险下有最大期望收益的点落在与马柯维茨弹头上面部分相切的直线上, 如图 10 所示.

称图 10 中的切线为资本市场线 (capital market line), 称马柯维茨有限前沿上的那个切点为市场组合 (market portfolio, 或者 capital market portfolio)

读者可以回忆一下我们前面关于马柯维茨曲线的平缓度的讨论. 由该讨论可知, 如果无风险利率过大, 将不存在资本市场线, 因此也不会有市场组合.

假定资本市场线存在, 那么我们可以通过调整投资组合中无风险资产和风险资产的比例, 也就是调整权重 w_{rf} 和 w_{risky}, 来实现资本市场线上的任一点. 为了实现市场组合右边的点, 需要卖空无风险资产, 然后用这些钱去买进更多的市场组合.

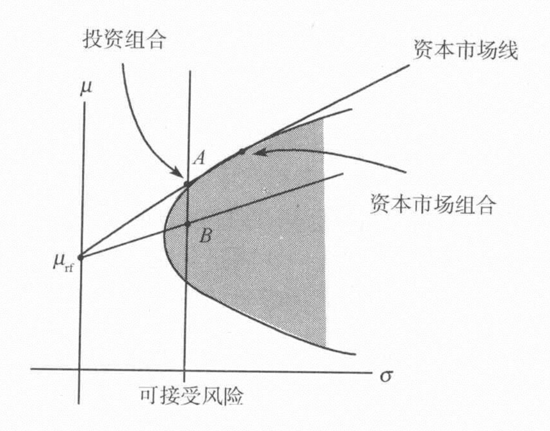


图 10　一给定风险水平下的投资组合

我们现在来表述上述讨论的要点：

在一给定风险水平下，为了使期望收益最大，投资者应该选择由无风险资产和市场组合 (无其他风险组合) 所组成的投资组合. 根据风险承受能力，相应地调整它们之间的比例.

资本市场线方程

假设市场组合的风险 / 期望收益点为 (σ_M, μ_M)，那么资产市场线的方程为

$$\mu = \mu_{\rm rf} + \frac{\mu_M - \mu_{\rm rf}}{\sigma_M}\sigma.$$

对资本市场线上的任意一点，都有

$$\mu - \mu_{\rm rf} = \frac{\mu_M - \mu_{\rm rf}}{\sigma_M}\sigma,$$

$\mu - \mu_{\rm rf}$ 为投资组合的期望收益超出无风险收益的额外部分，称其为风险溢价 (risk premium). 风险溢价是对风险的回报. 当然，风险的存在则暗示着投资者不一定真能得到这部分超额收益.

为了使资本市场线方程更具体，我们对市场组合的风险 / 期望收益 (σ_M, μ_M) 需要了解更多. 市场组合中的各资产权重表达式将由下面定理给出.

定理 6　*对任意无风险利率 $\mu_{\rm rf}$，资本市场组合的权重为*

$$W = \frac{(M - \mu_{\rm rf}O)C^{-1}}{(M - \mu_{\rm rf}O)C^{-1}O^{\rm T}}.$$

注意，分母是一个实数，为分子中的向量的各分量之和.

证明 对马柯维茨弹头里的任意一点 (σ,μ), 过点 $(0,\mu_{\mathrm{rf}})$ 和 (σ,μ) 的直线的斜率为

$$s=\frac{\mu-\mu_{\mathrm{rf}}}{\sigma}=\frac{\sum\mu_iw_i-\mu_{\mathrm{rf}}}{\sqrt{\sum c_{i,j}w_iw_j}}.$$

从直观上看显然有, 切点具有这样的性质, 即过该点的直线的斜率是马柯维茨弹头里所有的点 (σ,μ) 中最大的. 接下来我们在满足约束条件 $\sum w_i=1$ 下最大化目标. 再次利用拉格朗日乘子, 求下面函数的偏导数, 再令它们等于 0,

$$f=\frac{\sum_i\mu_iw_i-\mu_{\mathrm{rf}}}{\sqrt{\sum_{i,j}c_{i,j}w_iw_j}}+\lambda\Big(1-\sum_i w_i\Big).$$

我们将计算过程留作练习, 经处理得到

$$\frac{\partial f}{\partial w_k}=\frac{\mu_k\sum c_{i,j}w_iw_j-(\sum\mu_iw_i-\mu_{\mathrm{rf}})(\sum c_{i,k}w_i)}{(\sum c_{i,j}w_iw_j)^{3/2}}-\lambda=0.$$

整理得

$$(WCW^{\mathrm{T}})\mu_k-(MW^{\mathrm{T}}-\mu_{\mathrm{rf}})C_kW^{\mathrm{T}}=\lambda(WCW^{\mathrm{T}})^{3/2},$$

其中 C_k 为协方差矩阵 C 的第 k 行. 上面方程还可写成

$$\sigma^2\mu_k-(\mu-\mu_{\mathrm{rf}})C_kW^{\mathrm{T}}=\lambda\sigma^3.$$

因为对所有的 k 都成立, 故

$$\sigma^2M^{\mathrm{T}}-(\mu-\mu_{\mathrm{rf}})CW^{\mathrm{T}}=\lambda\sigma^3O^{\mathrm{T}}.$$

取转置并利用 $C^{\mathrm{T}}=C$, 可得

$$\sigma^2M-(\mu-\mu_{\mathrm{rf}})WC=\lambda\sigma^3O.$$

两边同时右乘 W^{T} 并利用 $OW^{\mathrm{T}}=1$, 可得

$$\sigma^2MW^{\mathrm{T}}-(\mu-\mu_{\mathrm{rf}})WCW^{\mathrm{T}}=\lambda\sigma^3,$$

或者

$$\sigma^2\mu-(\mu-\mu_{\mathrm{rf}})\sigma^2=\lambda\sigma^3.$$

因此

$$\lambda=\frac{\mu_{\mathrm{rf}}}{\sigma}.$$

现将 λ 代入先前的方程中去, 有

$$\sigma^2 M - (\mu - \mu_{\rm rf})WC = \mu_{\rm rf}\sigma^2 O.$$

可写成

$$\frac{\mu - \mu_{\rm rf}}{\sigma^2} W = (M - \mu_{\rm rf}O)C^{-1}.$$

两边同时右乘 $O^{\rm T}$, 并注意到有 $WO^{\rm T} = 1$, 故得到

$$\frac{\mu - \mu_{\rm rf}}{\sigma^2} = (M - \mu_{\rm rf}O)C^{-1}O^{\rm T}.$$

利用上面这个方程, 就可以得到

$$W = \frac{(M - \mu_{\rm rf}O)C^{-1}}{(M - \mu_{\rm rf}O)C^{-1}O^{\rm T}}.$$

为了举例说明, 我们继续例 1 来说明如何求得市场组合.

例 2 继续例 1, 图 11 显示了我们的 Excel 工作的另外部分. 求解了基于不同无风险利率下的市场组合的风险和期望收益 (这里只有 3 种利率).

市场组合							
无风险利率(RFR)	M-RFR*O			(M-RFR*O)C^{-1}			Den
0.01	0.09	0.1	0.06	1.883595808	1.659126341	0.433966462	3.976688611
0.02	0.08	0.09	0.05	1.701097806	1.509556481	0.29421513	3.504869417
0.03	0.07	0.08	0.04	1.518599804	1.359986621	0.154463797	3.033050222

W			收益 μ	WC			$\sigma_2=WCW^{\rm T}$	风险 σ
0.473659366	0.417213039	0.109127594	0.100898303	0.022631895	0.02514655	0.01508793	0.022857787	0.151187921
0.485352692	0.430702632	0.083944677	0.101788686	0.022825387	0.02567856	0.014265867	0.02333573	0.152760368
0.500684029	0.448389087	0.050926884	0.102956084	0.023079077	0.026376088	0.013188044	0.024053701	0.155092557

图 11

比如, 当无风险利率 $\mu_{\rm rf} = 0.03$ 时, 市场组合的期望收益和风险为

$$\mu_m = 1.02956084,$$
$$\sigma_m = 0.155092557.$$

对市场组合的进一步认识

根据我们的理论, 所有的理性投资者都会选择投资于市场组合和无风险资产上. 关于市场组合还有更深刻的结论. 首先, 市场组合必须包含所有可能的资产! 因为如果某一种资产不在该组合里, 那么没有人去买, 该资产就会衰退消亡.

既然市场组合包含了所有的资产, 那么该组合就没有非系统风险 —— 该类风险已经被完全分散化了. 因此, 市场组合中的所有风险均为系统风险.

在实践中, 可用数目较少的资产种类来逼近代替市场组合. 研究表明, 经细心筛选的 20 ~ 40 种证券组成的投资组合达到了一定程度的风险分散化和逼近真实的市场组合. 我们将用市场组合这个词语来指示一个未表明的风险高度分散化的投资组合, 因为是高度分散化的, 所以可看成是没有非系统风险的.

资产的风险-收益与市场组合的关系

我们考虑市场组合中任一特定资产 a_k. 利用在第 1 章中讨论的最佳线性估计, 通过市场组合收益函数来逼近资产 a_k 的收益 R_k. 根据第 1 章的定理 9, 我们可以写成

$$R_k = \beta_k R_M + \alpha_k + \varepsilon,$$

其中

$$\begin{aligned}\beta_k &= \frac{\text{cov}(R_k, R_M)}{\sigma_M^2},\\ \alpha_k &= \mathcal{E}(R_i) - \beta_k \mathcal{E}(R_M).\end{aligned}$$

而 ε 则为误差 (剩余随机变量). 系数 β_k 是资产收益相对于市场组合收益的贝塔 (beta) 值, 也是线性回归的斜率.

为了感受到我们期望做什么, 图 12 给出了在贝塔值相对大和误差有 3 种数量级 (从非常小到相当大) 情形下的最佳线性估计.

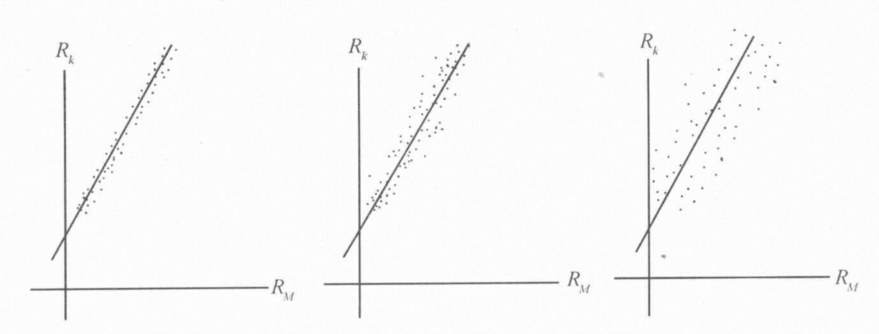


图 12 大贝塔值和不同数量级误差

因为贝塔值比较大, 所以在这三种情形下, 当市场收益变动一定数量时, 资产的收益变动了相对较大的数量. 从另一方面考虑, 如果市场收益在一定范围内波动 (例如用方差度量), 那么资产收益的波动相对较大 (用方差度量). 因此, 市场风险在资产风险中被“放大”.

图 13 给出了贝塔值比较小时的最佳线性估计, 误差也有 3 种不同数量级.

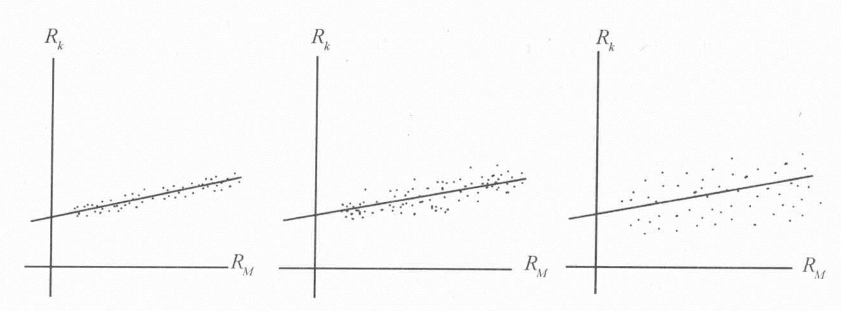


图 13 小贝塔值和不同数量级误差

因为贝塔值比较小, 所以在这三种情形下, 当市场收益变动一定数量时, 资产的收益变动了相对较小的数量. 因此, 市场风险在资产风险中被“缩小”.

直觉告诉我们, 一种资产的系统风险, 即来自于资产与市场组合 (它的风险全为系统风险) 之间的关系, 在某种方式上与资产的贝塔值有关.

另外, 从图 12 和 13 中可以看出, 还有另外一种因素影响着资产的风险, 该因素与市场风险没有关系. 这就是误差. 误差 ε 越大, 比如用方差 $\mathrm{var}(\varepsilon)$ 度量, 资产期望收益的不确定性就越大.

现在我们利用数学知识来看看我们验证上面的陈述. 事实上, 对于单个资产 R_k 的期望收益和风险, BLP 给出了其 β 形式的公式.

风险

对于风险来说, 我们有 (证明过程留作练习)

$$\mathrm{cov}(R_M, \varepsilon) = 0.$$

因此, 资产 a_k 的风险为 (因 α_k 为常数)

$$\begin{aligned}\sigma_k^2 = \mathrm{var}(R_k) &= \mathrm{var}(\beta_k R_M + \alpha_k + \varepsilon)\\ &= \mathrm{var}(\beta_k R_M + \varepsilon)\\ &= \beta_k^2 \sigma_M^2 + \mathrm{var}(\varepsilon).\end{aligned}$$

因此, 称 $\beta_k^2\sigma_M^2$ 为资产 a_k 的系统风险, 与市场风险成比例, 比例系数为 β_k^2. 资产风险中的剩余部分就是 $\mathrm{var}(\varepsilon)$, 为先前讨论的误差度量. 称它为资产的非系统风险或者独特风险 (unique risk).

根据经济学理论, 当向分散化的投资组合里加入一种资产时, 该资产的非系统性风险能被组合中的其他资产对冲掉. 因此, 当估计资产的风险 – 收益表现时, 非系统性风险不应被考虑在内. 故资产的贝塔值成为资产的风险 – 收益分析中的焦点.

期望收益

为了更深入地证明上述观点是正确的, 我们来考虑市场组合的期望收益

$$\mu_M = MW_M^{\mathrm{T}}$$

和单个资产 a_k 的期望收益

$$\mu_k = Me_k^{\mathrm{T}},$$

其中 $e_k = (0 \ \cdots \ 0 \ 1 \ 0 \ \cdots \ 0)$ 是单位矩阵的第 k 行. 为了使上述两个表达式有联系, 我们需要知道 M 的表达式.

回顾一下, 定理 6 给出了市场组合中各资产的权重表达式为

$$W = \frac{(M - \mu_{\mathrm{rf}}O)C^{-1}}{(M - \mu_{\mathrm{rf}}O)C^{-1}O^{\mathrm{T}}}.$$

既然分母为一常数, 我们就记它的倒数为 δ. 故

$$W = \delta(M - \mu_{\mathrm{rf}}O)C^{-1}.$$

解之得

$$M = \frac{1}{\delta}W_M C + \mu_{\mathrm{rf}}O.$$

现在我们就可将 μ_M 和 μ_k 写成下面的形式

$$\begin{aligned}\mu_M &= MW_M^{\mathrm{T}} = \left(\frac{1}{\delta}W_M C + \mu_{\mathrm{rf}}O\right)W_M^{\mathrm{T}} \\ &= \frac{1}{\delta}W_M C W_M^{\mathrm{T}} + \mu_{\mathrm{rf}}OW_M^{\mathrm{T}} \\ &= \frac{1}{\delta}\sigma_M^2 + \mu_{\mathrm{rf}}\end{aligned}$$

和

$$\begin{aligned}\mu_k &= Me_k^{\mathrm{T}} = \left(\frac{1}{\delta}W_M C + \mu_{\mathrm{rf}}O\right)e_k^{\mathrm{T}} \\ &= \frac{1}{\delta}W_M C e_k^{\mathrm{T}} \mu_{\mathrm{rf}} O e_k^{\mathrm{T}} \\ &= \frac{1}{\delta}\mathrm{cov}(R_k, R_M) + \mu_{\mathrm{rf}}.\end{aligned}$$

现在, 读者可能已经注意到上面这些式子中的一些部分和贝塔值之间的相似性, 贝塔值表达式如下

$$\beta_k = \frac{\mathrm{cov}(R_k, R_M)}{\sigma_M^2}.$$

通过前面的方程解出贝塔的分子和分母, 再代入到贝塔的表达式中去, 得

$$\beta_k = \frac{\mathrm{cov}(R_k, R_M)}{\sigma_M^2} = \frac{\delta(\mu_k - \mu_{\mathrm{rf}})}{\delta(\mu_M - \mu_{\mathrm{rf}})} = \frac{\mu_k - \mu_{\mathrm{rf}}}{\mu_M - \mu_{\mathrm{rf}}}.$$

最后解 μ_k 得

$$\mu_k = \mu_{\mathrm{rf}} + \beta_k(\mu_M - \mu_{\mathrm{rf}}).$$

让我们来把这些重要的公式集中放到一个定理中.

定理 7 市场组合中的资产 a_k 的期望收益和方差与该资产相对于市场组合的贝塔值有如下关系

$$\begin{aligned}\mu_k &= \mu_{\mathrm{rf}} + \beta_k(\mu_M - \mu_{\mathrm{rf}}),\\ \sigma^2 &= \beta_k^2\sigma_M^2 + \mathrm{var}(\varepsilon),\end{aligned} \tag{2.3.1}$$

其中 ε 是误差 (剩余随机变量).

预期收益的表达式证实了我们先前讨论的结果: 资产的期望收益只依赖资产的系统风险 $\beta_k^2\sigma_M^2$ (通过贝塔值), 而不依赖它的非系统风险 $\mathrm{var}(\varepsilon)$. 也证明了当评价资产相对市场组合的风险时只需考虑 $\beta_k^2\sigma_M^2$ 这一项.

因为在正常条件下, $\mu_M - \mu_{\mathrm{rf}}$ 为正, 故线性关系中的斜率为正, 这意味着贝塔值越大期望收益也越大, 反之也一样. 这是合理的 —— 在市场均衡下, 系统风险越大的资产要求的期望收益也应该越高.

当然, 没有规则说越高的风险应该得到越高的期望收益. 但是, 这是市场达到均衡的条件. 如果一种资产的收益少于它在市场上本来应有的收益, 那么没有人会买这种资产, 它的价格会自然下降, 但是这就提高该资产的收益. 同样地, 如果资产的收益比根据它的风险水平而在市场上应有的收益来得高的话, 那么会有更多的投资者买它, 因而会抬高它的价格而降低预期收益.

利用数学知识就能证实这点. 例如, 如果一种资产的系统风险低于市场组合的风险, 也就是说, 如果

$$\beta_k^2\sigma_M^2 < \sigma_M^2,$$

或者等价地有 $\beta_k < 1$, 那么该资产的收益满足

$$\mu_k = \mu_{\mathrm{rf}} + \beta_k(\mu_M - \mu_{\mathrm{rf}}) < \mu_{\mathrm{rf}} + (\mu_M - \mu_{\mathrm{rf}}) = \mu_M.$$

这说明了该资产的期望收益小于市场组合的期望收益. 另一方面, $\beta_k = 1$ 时有 $\mu_k = \mu_M$; $\beta_k > 1$ 时有 $\mu_k > \mu_M$. 这正如在市场均衡时我们所希望的那样.

称由定理 7 中第一个方程所描述的直线为证券市场线 (security market line, SML). 该方程指出了资产的期望收益等于无风险收益加上资该产的风险溢价 $\beta_k(\mu_M - \mu_{\mathrm{rf}})$.

例 3 假设无风险收益为 3%, 市场组合的期望收益为 12%. 考虑如下资产和它们的贝塔值.

资产	a_1	a_2	a_3	a_4	a_5
贝塔系数	0.65	1.00	1.20	−0.20	−0.60

因为

$$\mu_M - \mu_{\mathrm{rf}} = 0.12 - 0.03 = 0.09.$$

所以证券市场线方程为

$$\mu_k = 0.09\beta + 0.03.$$

我们现在可以计算在市场均衡下的期望收益.

资产	a_1	a_2	a_3	a_4	a_5
贝塔系数	0.65	1.00	1.20	−0.20	−0.60
期望收益 μ_k	8.85	12	13.8	1.2	−6.9

上面表格中的期望收益值是市场根据市场组合的全部系统风险 (和无风险利率) 提供的收益值. 例如, 资产 a_1 的贝塔值小于 1, 它的系统风险低于市场组合的系统风险, 因此市场给它提供的期望收益低于市场组合的期望收益 $8.85\% < 12\%$. 资产 a_2 与市场组合有相同的系统风险, 故市场给它提供的期望收益与市场组合的一样.

练 习 2

1. 市场组合的贝塔值是多少? 一个投资组合能将任何实数作为自己的贝塔值吗?

2. 设无风险利率为 4%, 市场组合的期望收益为 8%. 写出证券市场线方程.

3. 证明下面的参数方程组在 (σ', μ) 平面上表示一条直线

$$\mu = (1-s)\mu_1 + s\mu_2,$$
$$\sigma' = (1-s)\sigma_1 \pm s\sigma_2,$$

这里 s 为参数, 取遍所有的实数. 先考虑正号, 然后考虑负号.

4. 当 $\rho_{1,2} = 1$ 时我们有 $R_2 = aR_1 + b$, $a > 0$. 计算风险为 0 时的期望收益.

5. 证明在假设 $\sigma_1 \leqslant \sigma_2$和$\rho_{1,2} < 1$ 下有

$$s_{\min} = \frac{\sigma_1(\sigma_1 - \rho_{1,2}\sigma_2)}{\sigma_1^2 + \sigma_2^2 - 2\rho_{1,2}\sigma_1\sigma_2} < 1.$$

6. 设 ε 为资产 a_i 关于市场组合的最佳线性估计的误差, 试证明 $\mathrm{cov}(R_M, \varepsilon) = 0$.

7. 证明市场组合中的所有资产的回归线经过同一点. 该点是哪一点?

8. 令 P_M 为市场组合, 所包含的资产 a_i 的权重为 w_i. 将 R_i 的最佳线性估计写成

$$\mathrm{BLP}(R_i) = \beta_i R_M + \alpha_i.$$

考虑前两种资产 a_1和a_2, 其相应权重为 w_1 和 w_2. 这两种资产的收益和为

$$R_0 = w_1 R_1 + w_2 R_2.$$

如果最佳线性估计为

$$\mathrm{BLP}(R_0) = \beta_0 R_M + \alpha_M.$$

请问 β_0, β_1和β_2 之间有何关系? α_1, α_2和α_3 之间呢? 对市场组合中的任何资产子集(即任意子组合) 能得到相同的结论吗?

9. 检验一下图 14 中的数据 (至少在一些小数点的地方). 请说明为了达到 5% 的收益, 承受的风险至少为 0.238952. 如果无风险收益为 3%, 请证明市场组合的权重为 (0.335, 0.372, 0.293), 其风险 – 收益为 (0.194, 0.196).

资本资产定价模型: 灰色单元为输入区

用户数据		收益 μ_i	风险 σ_i	相关性	
	i=1	0.1	0.2	−0.1	$=\rho_{1,2}=\rho_{2,1}$
	i=2	0.2	0.3	0.2	$=\rho_{1,3}=\rho_{3,1}$
	i=3	0.3	0.4	0.2	$=\rho_{2,3}=\rho_{3,2}$
$C=(c_{i,j})=(\rho_{i,j}\sigma_i\sigma_j)$		j=1	j=2	j=3	
	i=1	0.04	−0.006	0.016	
	i=2	−0.006	0.09	0.024	
	i=3	0.016	0.024	0.16	
C的逆		26.6075388	2.58684405	−3.048780488	
		2.58684405	11.8255728	−2.032520325	
		−3.048780488	−2.032520325	6.859756098	
最小风险点		OC^{-1}=	26.14560237	12.37989653	1.778455285
		$OC^{-1}O^{T}$=	40.30395418		
		$W=OC^{-1}/OC^{-1}O^{T}$=	0.648710602	0.307163324	0.044126074
		$\mu=MW^{T}$=	0.139541547		
		WC=	0.024811461	0.024811461	0.024811461
		$\sigma^2=MCW^{T}$=	0.024811461		
		σ=	0.157516543		
最小风险线		MC^{-1}=	2.263488544	2.014042868	1.346544715
		$MC^{-1}O^{T}$=	5.624076127		
		$OC^{-1}M^{T}$=	5.624076127		
		$MC^{-1}M^{T}$=	1.033120843		
		Denom Det=	10.00862281		

图 14

第 3 章　期权的背景知识

在研究衍生品定价模型之前, 我们需要先了解股票期权的基础知识. 熟悉这些衍生品的读者可跳过该章内容, 而只需了解术语, 就像过去一样.

3.1　股票期权

股票期权有两类: 卖权 (putoptions) 和买权 (calloptions). 它们的定义如下.

定义 1　*买权就是它的卖方 (writer or seller) 和买方 (buyer) 之间的一个合约. 买方有权以一固定价格 (该价格称为执行价, strike price 或者 exercise price, 经常用 K 或者 E 来表示) 从卖方手里买入股票. 欧式买权 (European call) 只能在期权的到期日才能执行. 美式买权 (American call) 可在到期日前的任意时间执行.*

卖权就是它的卖方 (writer or seller) 和买方 (buyer) 之间的一个合约. 买方有权以一固定价格 (该价格称为执行价, strike price 或者 exercise price) 将股票卖给卖方. 欧式卖权 (European put) 只能在期权的到期日才能执行. 美式卖权 (American put) 可在到期日前的任意时间执行.

期权的卖方为空头方, 买方为多头方.

3.2　期权的用途

最初, 期权是为了用来对冲 (hedging) 和投机 (speculation) 的 (如果可能的话, 套利总是有好处的). 同时, 期权相对于拥有标的资产具有一个很重要的优点, 即杠杆效应 (leverage).

对冲是一种能降低现有头寸风险的投资. 为了通过例子说明期权的对冲特性, 假设一个投资者目前 (10 月份) 拥有 1000 股 IBM 的股票, 当前股价为 \$88 每股. 该投资者很关心的是, 如果股价大跌, 将给他的投资造成重大的打击.

因此该投资者可购买 12 月份执行价为 \$85 的卖权来对冲股价下跌. 这样, 他在接下来的 3 个月可以以每股 \$85 的价格卖出股票. 因此, 如果股价下跌严重, 投资者可以每股 \$85 的价格抛出. 对冲所需的代价为卖权的价格, 该价格当前为 \$1.5. 故只需花费 \$1500 就能保护 \$88000 资产.

杠杆效应

IBM 现在的价格为每股 \$90, 一个资金量为 \$450 的小型投资者只能购买 5 股 IBM 股票. 如果该投资者认为股价将会明显上涨, 那么利用期权可以使他避免资金量不足而投机该股票, 这比直接购买股票更有效.

例如, 期限为 1 个月执行价为 \$90 的买权当前价格为 \$3.80. 因此, 投资者可以购买 118 份这样的买权 (忽略佣金等). 如果在执行时间股价上涨到 \$95, 5 股股票产生的利润为 \$25, 而购买买权产生的利润为 \$590. 投资期权的收益率超过了 100%, 而直接投资股票的收益率还不到 6%. 这就是杠杆效应.

当然, 买权也有不利的一面, 如果股价没有上涨或者在到期日前没有上涨, 那么投资者将不能从期权上得到一分钱, 还需花费购买期权所需的佣金等, 而股票持有者可继续拥有该只股票.

3.3 利润曲线和损益曲线

当到期日来临时, 当且仅当有正的收益, 期权的持有者才会执行期权. 因此, 如果期权的执行价为 K, 股票的现货价格为 S, 那么买权的持有者将会执行该期权当且仅当 $K < S$. 另一方面, 卖权的持有者将会执行该期权当且仅当 $K > S$. 一些术语被用来描述不同的可能性.

定义 2 称买权

1) 处于实值状态(in the money), 如果 $K < S_T$.

2) 处于两平状态(at the money), 如果 $K = S_T$.

3) 处于虚值状态(out of the money), 如果 $K > S_T$.

称卖权

1) 处于实值状态 (in the money), 如果 $K > S_T$.

2) 处于两平状态 (at the money), 如果 $K = S_T$.

3) 处于虚值状态 (out of the money), 如果 $K < S_T$.

重要的一点是, 一个期权处于实值状态并不意味着其拥有者能够获得利润. 问题是初始成本 (还有委托等费用, 在整个讨论中我们都不考虑这些) 可能超出执行期权所得到的收入. 在这种情况下, 投资者还是会执行期权, 这是因为正的收益能够帮助降低总损失.

图 1 给出了期权每种头寸下的损益 (pay off)(忽略成本). 注意横轴表示到期时的股价, 所有的线段部分要么在横轴上要么斜率就为 ± 1.

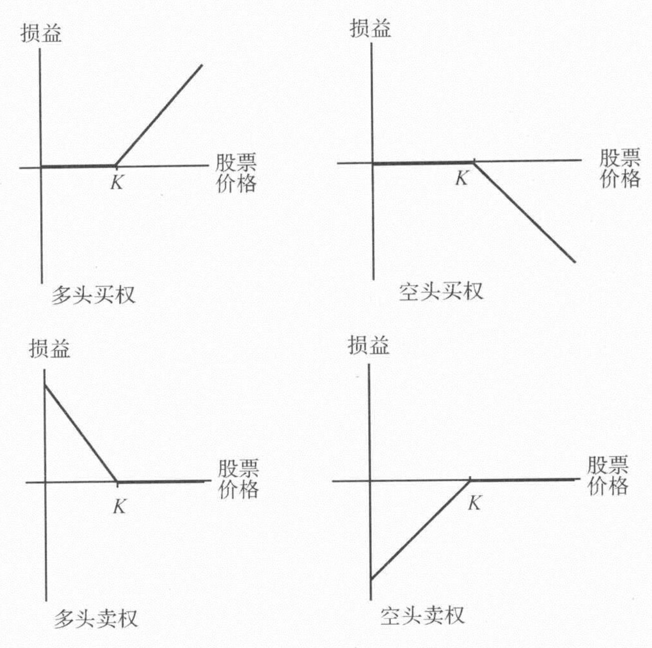


图 1　损益曲线

例如, 买权的多头方会执行期权当且仅当股票的现货价格 S 大于 K. 在该情形下, 持有者的收入为 $S-K$. 否则, 持有者会让买权到期而失效.

图 2 给出了真实利润曲线, 在这里面考虑了购买或者出售期权的成本 (正如提过的, 我们忽略其他交易费).

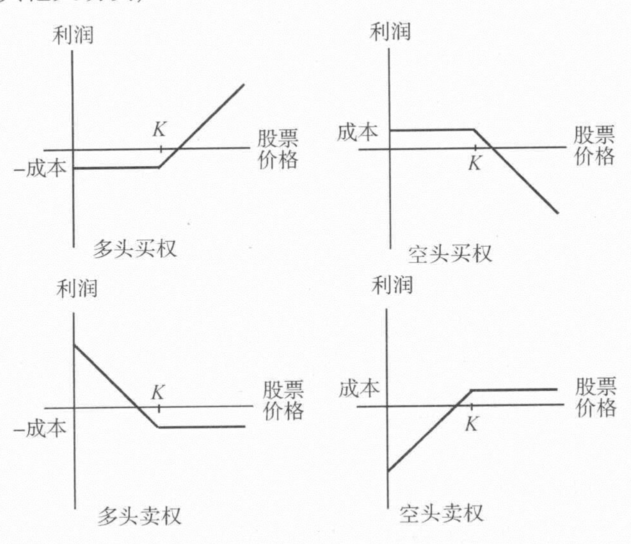


图 2　利润曲线

损益公式实际上很简单. 对一个买权的多头来说, 如果股价 S 满足 $S \geqslant K$, 那么执行买权后得到的收入为 $S-K$; 而当 $S<K$ 时, 买权到期后失效, 收入为 0. 因此有

$$\text{payoff}(\text{多头买权}) = \max\{S-K, 0\}.$$

对卖权来说, 有

$$\text{payoff}(\text{多头卖权}) = \max\{K-S, 0\}.$$

如所提及的那样, 损益曲线是非常有用的. 从这些曲线中我们能立即得到下面一些结论.

多头买权:

- 有下界: 下界为买权的成本.
- 无上界: 因为股价没有上限.
- 乐观的头寸: 买方希望股价上涨.
- 当股价大于执行价 K 时, 应该执行多头买权.

空头买权:

- 无下界: 因为股价没有上限.
- 有上界: 上界为买权的售价.
- 悲观的头寸: 买方希望股价下跌.
- 当股价下跌到小于执行价 K 时, 应该执行空头买权.

多头卖权:

- 有下界: 下界为卖权的成本.
- 有上界: 因为股价最多只能下跌到 0, 这时的利润等于执行价乘以持有份数再减去卖权的成本.
- 悲观的头寸: 买方希望股价下跌.
- 当股价下跌到小于执行价时, 应执行多头卖权.

空头卖权:

- 有下界: 因为股价至多下跌到 0, 这时的损失等于执行价乘以持有份数再加上卖权的成本.
- 有上界: 上界为卖权的出售价格.
- 乐观的头寸: 买方希望股价上涨.
- 当股价上涨到大于执行价时, 应执行空头卖权.

值得注意的是, 多头买权和空头卖权都是牛市头寸 (bullish position), 即股价上涨时获利. 不过两者的区分是风险程度和获利程度. 同样地, 多头卖权和空头买权都是熊市头寸 (bearish position).

利润曲线也提供了一些有趣的信息. 先不考虑风险因素, 我们有下面结论：

• 如果我们相信股价会下跌, 但只是轻微的. 令股价在区间 $(K - \text{成本}, K]$ 内, 那么空头买权是最有优势的头寸. 在该情形下, 根据卖权的成本, 多头卖权仍不能获得利润. 但是, 如果我们觉得股价会大幅度下跌, 那么多头卖权最具优势.

• 如果我们相信股价会上涨, 但只是轻微的, 那么空头卖权是最有优势的头寸. 但是, 如果我们觉得股价会大幅度上涨, 那么多头买权最具优势.

有保护看涨期权 (covered calls)

我们已经说过空头买权的收益没有下界, 因为理论上股价能上涨到无穷大. 当然, 这假定了买权的卖方在到期时买进股票, 然后交割给买权的买方. 但是, 如果卖方已经拥有了股票, 那么他损失的下限就只是这些股票的价值, 因为他能用这些股票去"覆盖"买权.

如果买权的卖方在出售买权的同时拥有标的股票, 那么说他写了一张有保护的看涨期权. 写有保护的看涨期权比没有写保护的看涨期权 (也称裸期权,naked option) 安全很多.

期权组合的利润曲线

期权组合包含一系列期权. 下面的例子告诉了如何得到一个简单组合的利润曲线.

例 1 假设所有期权的到期日相同, 期权的买卖情况由下式给出

$$-P_{100} + P_{120} + 2C_{150} - C_{180},$$

即头寸情况是这样的：一张执行价为 100 的卖权空头, 一张执行价为 120 的卖权多头, 两张执行价为 150 的买权多头和一张执行价为 180 的买权空头. 在同一图上画出单个期权的利润曲线, 然后就能利用这些曲线就得到总利润曲线, 如图 3 所示. 注

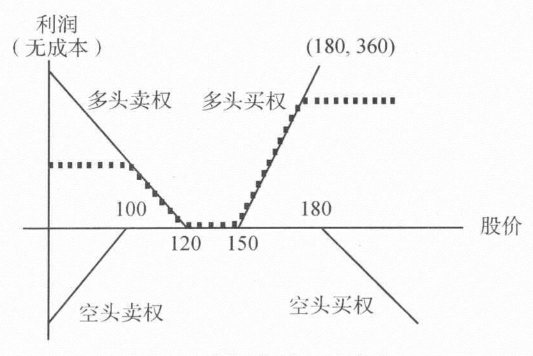


图 3 利润曲线 (无成本)

意到, 在绘制曲线时, 如果先忽略所有的成本会较简单. 然后简单地上下调动曲线一定值, 该值为涉及到的期权的总成本, 在该例子中为

$$-\mathrm{cost}(P_{100}) + \mathrm{cost}(P_{120}) + 2\mathrm{cost}(C_{150}) - \mathrm{cost}(C_{180}).$$

3.4 卖 空

如果对股票的卖空这个概念不了解的话, 那么对期权就不能进行完全分析. 为了卖空股票, 投资先借入股票 (一般从经纪人那里), 然后立即卖掉, 从而募集到的资金为股票的当前价格. 对于这种特别待遇, 投资者必须到时归还股票 (而不是钱) 给借出者.

与空头买权一样, 卖空股票会导致潜在的无下界的损失, 除非卖者也拥有这些股票, 因为可利用这些股票来对冲借入的股票的收益. 图 4 给出了卖空股票的利润曲线和股票多头的利润曲线图.

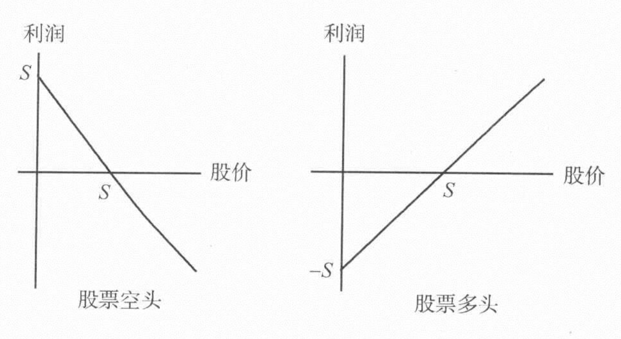


图 4 股票空头和多头利润曲线

练 习 3

1. 请不看书画出多头卖权、空头卖权、多头买权和空头买权的利润曲线图.

2. 为了写一张有保护的看跌期权, 投资者必须在写一张卖权的同时卖空相同数量的标的股票. 例如, 假设一个投资者写一张交割 100 股 IBM 股票的卖权, 同时卖空 100 股 IBM 股票. 意思是投资者借入 100 股 IBM 股票卖掉. 描述写这张有保护的看跌期权的上限和下限. 为什么有人会写有保护的看跌期权?

3. 画出下面期权组合的收益图

$$-P_{80} + P_{100} + 2C_{130} - C_{150}.$$

套利 (spread, 价差) 是指投资者同时买入和卖出不同种类的期权, 不过两种期权的标的物相同, 只是其他一些条款不同 (执行价或到期日). 看涨期权套利 (call

spread) 涉及买卖买权, 看跌期权套利 (put spread) 涉及买卖卖权. 这里的想法就是用一个期权去对冲另一期权的风险.

4. 在牛市套利中 (bull spread), 投资者买入执行价为 K_1 的买权, 卖出执行价 K_2 更高的买权, 两种买权的到期日一样. 画出牛市套利的利润曲线图. 牛市套利什么时候最有利可图? 是积极还是消极投资? 提示：首先必须确定如何比较两种买权的成本.

5. 在熊市套利中 (bear spread), 投资者买入执行价为 K_1 的买权, 卖出执行价 K_2 更低的买权, 两种买权的到期日一样. 画出熊市套利的利润曲线图. 熊市套利什么时候最有利可图? 是积极还是消极投资? 提示：首先必须确定如何比较两种买权的成本.

6. 在时间套利中 (calender spread or time spread), 投资者出售到期日为 D_1 的买权, 买入另一到期日 D_2 更远 (意思是 $D_2 \geqslant D_1$) 的买权. 假定两种买权的执行价相同. 考虑下面的时间套利. XYZ 的当前 (1 月) 价格为 \$50. 买权的价格如下：

APR 50 call：(到期日为 4 月执行价为 \$50)\$5,

JUL 50 call：\$8,

OCT 50 call：\$10.

假设 4 月份的股价仍是 \$50. 如果其他条件都不变, 那么期权的价格应该如下：

APR 50 call：(到期了)\$0,

JUL 50 call：\$5,

OCT 50 call：\$8.

投资者获利吗? 描述一下原因.

7. 蝶式套利 (butterfly spread) 是牛市套利和熊市套利的组合. 看涨蝶式套利买入一份执行价为 K_1 的买权, 卖出两份执行价为 K_2(大于K_1) 的买权, 同时卖出一份另一执行价为 K_3(大于K_2) 的买权. 所有的买权具有相同的到期日. 画出蝶式套利的利润曲线图.

第4章　套　　利

本章先简单介绍套利概念以及如何运用无套利假设给资产定价, 然后简要地讨论远期合约的定价和与期权定价相关的一些简单话题.

4.1　远期合约的背景知识

我们从远期合约的必要背景知识出发.

远期合约

远期合约是一张在将来某确定时间 (交割日) 按确定的数量确定的价格 K(交割价格) 购买某种资产 (标的资产) 的协议. 参与远期合约不需要付出任何初始价钱 —— 它是免费的.

同意买入资产的一方持有合约的多头, 称为合约的买方. 同意出售资产的一方持有合约的空头, 称为合约的卖方.

期货合约

与上面描述的香草型远期合约形成对比的是期货合约, 期货合约是有一系列约束条件和更复杂的收益模型的远期合约. 确实, 期货合约很少能持续至到期日, 也就是说, 在所有期货合约中只有非常小的一部分 (大概 1% ~ 2%) 能够生存到交割日. 期货合约的主要性质如下：

1) 期货合约在有组织的交易所里进行交易. 例如, 芝加哥期货交易所 (CBT 或 CBOT) 是最大的期货交易所.

2) 期货合约有标准化的条款, 如交割数量、标的物的精确类型、交割日和交割价格. 就如你在五金店只能买特定长度和直径的螺钉一样, 你只能买具有标准化条款的期货合约.

3) 期货合约的盈亏由清算所 (clearinghouse) 结算.

4) 购买期货合约需要买方提供保证金 (margin), 即需要一部分钱来支付每天潜在的亏损.

5) 期货市场由政府来调控, 而远期合约很大程度上没有调控.

6) 期货合约可通过交割、抵消 (即通过反向交易) 或者物理交易 (早期交割安排的一种形式) 来终止.

在本书中我们不具体讨论期货合约的细节.

远期价格

考虑一给定标的物 (比如小麦) 和交割日 T(比如 2003 年 12 月) 的远期合约. 在任一时间 $t < T$, 大家都可以交易该合约. 当然, 根据合约信息, 交割价格依赖时间 t, 记为 $F_{t,T}$. 称该价格为合约的远期价格.

例如, 一个提供在 9 月份交割 5000 蒲耳小麦的合约在 7 月 1 日的远期价格可能为 170 美分 (每蒲耳). 1 周后, 远期价格可能为 168 美分.

现货价格

与远期价格形成对比的是现货价格, 一种资产在时刻 t 的现货价格 S_t 就是该资产在时刻 t 立即交割的价格. 例如, 我们能谈及 1 蒲耳小麦的当前现货价格, 我们也能谈及未来 1 个月小麦的现货价格. 这个价格就是在未来 1 个月的这个时候立即交割, 投资者愿意支付的价格. 当然, 在当前时候, 这个现货价格是不知道的.

4.2 远期合约的定价

为了确定远期合约的远期价格, 我们运用一个简单的无套利原理. 假设远期合约标的资产每份的初始价格为 S_0(比如一张小麦合约的每份标的资产为 5000 蒲耳小麦). 考虑下面的两个投资组合.

组合 A：多头合约,

一张远期合约;

组合 B：现金和标的资产,

一份标的资产和 S_0 美元的负债.

在完美市场上, 能够对组合 A 或者组合 B 进行多头或空头. 为了空头组合 A, 我们卖出远期合约. 为了空头组合 B, 我们借入资产并以价格 S_0 卖出, 然后将所得资金借出. 这被称为反向 – 买现卖期 (reverse cash-and-carry) 策略.

这些组合的初始价值为 0. 最终的损益为

$$\begin{aligned} V(\text{多头合约}) &= S_T - F_{0,T}, \\ V(\text{空头合约}) &= F_{0,T} - S_T, \\ V(\text{买现卖期}) &= S_T - S_0 \mathrm{e}^{rT}, \\ V(\text{反向 – 买现卖期}) &= S_0 \mathrm{e}^{rT} - S_T. \end{aligned}$$

例如, 在买现卖期情形, 到时刻 T 时, 投资者拥有价值 S_T 的资产, 但是需要偿还 $S_0\mathrm{e}^{rT}$ 这么多的贷款.

考虑如下的两种策略.

策略 1. 多头合约和反向 – 买现卖期资产. 该策略的最终损益为

$$V(\text{多头合约}) + V(\text{反向 – 买现卖期}) = S_0\mathrm{e}^{rT} - F_{0,T}.$$

策略 2. 空头合约和正向 – 买现卖期资产. 该策略的最终损益为

$$V(\text{空头合约}) + V(\text{正向 – 买现卖期}) = F_{0,T} - S_0\mathrm{e}^{rT}.$$

如果两种策略中的任一策略损益为正, 那么投资者就有套利策略. 因此, 根据无套利原理有

$$S_0\mathrm{e}^{rT} - F_{0,T} = 0,$$

即

$$F_{0,T} = S_0\mathrm{e}^{rT}.$$

定理 1 *考虑这样的一个远期合约, 买方在未来 T 时刻买进现价为 S_0 的标的资产. 在完美无套利市场上, 远期价格为*

$$F_{0,T} = S_0\mathrm{e}^{rT},$$

其中 r 为无风险利率.

4.3 买权和卖权的平价公式

我们也能用无套利原理来推出标的物相同, 条款相同 (执行价和到期日) 的买权和卖权价格之间的关系.

首先, 我们对无风险债券做一解释. 我们假设能够买卖任意数量的无风险债券 (比如美国国库券). 债券的连续复合利率为 r, 并假设一份无风险债券在 0 时刻的价值为 1 美元.

非常重要的一点是我们在脑子里应该对债券的数量和价格保持区分. 例如, 假设我们在 0 时刻投资 A 份无风险债券, 则到 t 时刻时债券的数量仍是 A, 但是价值为 $A\mathrm{e}^{rt}$. 同样, 如果我们在 t 时刻投资 K 美元于无风险债券, 那么我们得到 $K\mathrm{e}^{-rt}$ 份债券, 每份债券价值 e^{rt}.

使用下面的记号会显得很方便:

$$X^+ = \max\{X, 0\}.$$

欧式情形

买权和卖权的平价公式比较了具有相同标的股票相同到期日和执行价 K 的欧式卖权价格 P 和欧式买权价格 C.

假设标的股票不分红且目前价格为 S_0. 考虑如下两个投资组合.

组合 A: 一份卖权多头, 一份买权空头. 该组合的初始价值为 $P-C$, T 时刻的损益为

$$\max(K-S_T,O)-\max(S_T-K,0)=K-S_T,$$

其中 S_T 是股票在 T 时刻的价格.

组合 B: 一份股票空头, 拥有价值为 $K\mathrm{e}^{-rT}$ 的无风险债券. 该组合的初始价值为

$$K\mathrm{e}^{-rT}-S_0.$$

最终损益为

$$K-S_T.$$

既然两个组合的最终损益一样, 那么初始价值必须也一样. 否则, 投资者就可以卖出较贵的组合而买进较便宜的组合, 这样在 T 时刻就能保证获得利润. 因此

$$P-C=K\mathrm{e}^{-rT}-S_0.$$

这就是买权和卖权的平价公式.

如果股票分红, 那么分析就有一点点不同. 原因是在组合 B 情形下, 股票借出者到时不但要求收回自己的股票, 而且还会索取股票在这期间的红利. 假设红利折现到时刻 0 的价值为 d_0, 则组合 B 的最终损益为

$$K-S_T-d_0\mathrm{e}^{rT}.$$

因此, 我们不能比较两个组合的初始价值, 因为两组合的最终损益不相等. 这就需要我们调整组合 B 使得其损益与组合 A 一样.

组合 B′: 一份股票空头, 拥有价值为 $K\mathrm{e}^{-rT}+d_0$ 的无风险债券. 该组合的初始价值为

$$K\mathrm{e}^{-rT}+d_0-S_0.$$

最终损益为

$$K+d_0\mathrm{e}^{rT}-S_T-d_0\mathrm{e}^{rT}=K-S_T.$$

现在我们就能将组合 A 和组合 B′ 的初始价值等同起来, 故

$$P - C = K\mathrm{e}^{-rT} + d_0 - S_0.$$

定理 2(带分红的欧式期权) 假设标的股票目前的价格为每股 S_0, 欧式卖权的价格为 P, 欧式买权为 C, 两者有相同的执行价 K 和到期日 T. 并假定在我们考虑期间内每股股票产生的红利的目前价值为 d_0. 那么在无套利假设下, 有

$$C - P = S_0 - K\mathrm{e}^{-rT} - d_0,$$

这里 r 为无风险利率. 称该公式为买权和卖权的平价公式.

美式情形

美式期权的情形比欧式情形复杂. 这里的价差 $C - P$ 不是一个常数, 虽然在欧式情形下为一常数.

考虑如下策略：买入一份卖权, 卖空一份买权, 买入一股标的股票 (为了应对买权被执行). 因此初始组合为：

1) 1 份卖权 @ P.

2) -1 份买权 @ C.

3) 1 股标的股票 @ S_0.

组合的初始价值为

$$V_0 = P - C + S_0.$$

在持有期内, 下列情形之一会发生：买权被执行或者失效. 如果买权在时刻 t 被执行, 那么我们卖出股票, 得到现金 $K\mathrm{e}^{-rt}$. 因此到最后时刻 T 时我们拥有的头寸如下：

1) 1 份卖权 @ $(K - S_T)^+$.

2) $K\mathrm{e}^{-rt}$ 美元 @ e^{rt}.

3) $d_0\mathrm{e}^{rT}$ 美元, 如果买权在分红后被执行.

故最后的价值为

$$V_{T,1} = (K - S_T)^+ + K\mathrm{e}^{r(T-t)} + d_0\mathrm{e}^{rT}\delta.$$

如果买权在分红后被执行, 则 $\delta = 1$, 否则 $\delta = 0$.

如果买权没有被执行, 我们一直将组合持有到最后, 这时的持有情况为：

1) 1 份卖权 @ $(K - S_T)^+$.

2) -1 份买权 @ 0.

3) 1 股标的股票 @ S_T.

4) $d_0\mathrm{e}^{rT}$ 美元.

最后的价值为

$$V_{T,2} = (K - S_T)^+ + S_T + d_0\mathrm{e}^{rT} = \max\{S_T, K\} + d_0\mathrm{e}^{rT}.$$

上述两个最终价值表达式有一点点复杂, 不过我们能得到其上下界. 实际上, $V_{T,1}$ 和 $V_{T,2}$ 满足

$$K \leqslant V_{T,1} \leqslant K\mathrm{e}^{rT} + d_0\mathrm{e}^{rT},$$
$$K \leqslant V_{T,2} \leqslant K\mathrm{e}^{rT} + d_0\mathrm{e}^{rT}.$$

对于该组合, 一个选择就是初始投资 V_0 在无风险债券上, 最后得到的损益为

$$V_0\mathrm{e}^{rT} = (P - C + S_0)\mathrm{e}^{rT}.$$

现在如果 $V_0\mathrm{e}^{rT} < K$, 那么债券投资策略得到的收益少, 这就存在无套利了：卖出债券并买进组合. 这种操作的初始成本为 0, 而最终损益至少

$$K - V_0\mathrm{e}^{rT} > 0.$$

因此, 为了避免有套利机会, 必须有

$$K \leqslant (P - C + S_0)\mathrm{e}^{rT}.$$

稍微懂一点代数知识就知道

$$C - P \leqslant S_0 - K\mathrm{e}^{-rT}.$$

同样地, 如果 $V_0\mathrm{e}^{rT} > K\mathrm{e}^{rT} + d_0\mathrm{e}^{rT}$, 那么债券投资策略得到的收益就比组合高, 因此这也存在套利机会：买债券并卖组合. 故为了避免存在套利机会, 必须有

$$(P - C + S_0)\mathrm{e}^{rT} \leqslant K\mathrm{e}^{rT} + d_0\mathrm{e}^{rT},$$

即有

$$S_0 - K - d_0 \leqslant C - P.$$

定理 3(带分红的美式期权)　假设标的股票目前价格为每股 S_0, 美式卖权的价格为 P, 美式买权为 C, 两者有相同的执行价 K 和到期日 T. 并假定在我们考虑期间内每股股票产生的红利的目前价值为 d_0. 那么在无套利假设下, 有

$$S_0 - K - d_0 \leqslant C - P \leqslant S_0 - K\mathrm{e}^{-rT},$$

其中 r 为无风险利率. 因此, 美式情形下的 $C - P$ 不会比欧式情形下的大, 但是可以比欧式情形下的小.

4.4 期权价格

结合平常的感觉,我们能从简单的套利讨论中得到一些有关期权价格的信息.例如,既然美式期权具有相应欧式期权的所有特征而且还更多,看起来似乎有美式期权不应该比欧式期权便宜. 用记号表示,即有

$$C^A \geqslant C^E, \quad P^A \geqslant P^E$$

我们将该问题留给读者,让读者给出一个任意的论据来支持上面的不等式.

不难明白,美式卖权的价格可能超过相应欧式卖权的价格. 想法是这样的,美式卖权能够提前执行,执行后能将一股股票转变为债券 (具有无风险收益 r). 如果无风险收益充分高,得到的利润可能比欧式卖权的高 (受执行价的限制,对欧式卖权买方最有利的情形也就是股价在 T 时刻为 0).

更具体地,假设债券的收益率为 r,股价为 S_0. 考虑一个执行价为 K 的美式卖权. 一方面,当股价在 T 时下跌到 0 时,相似的欧式卖权的最大收益为 K. 另一方面,假设我们在 0 时刻执行美式卖权,将所得收益 $K-S_0$ 投资在债券上. 最后得到的收益为 $(K-S_0)(1+r)$. 因此,如果有

$$(K-S_0)(1+r) > K.$$

那么,美式卖权比相似的欧式卖权更值钱.

现给出一数值例子,假设有一段时间债券的收益率为 12%(或者更高),股票的价格为 \$5. 考虑一个执行价为 \$50 的美式卖权,最终的收益

$$(K-S_0)(1+r) = 45 \times (1.12) = 50.40 > 50.$$

因此,执行价为 \$50 的美式卖权比执行价为 \$50 的欧式卖权更值钱.

另一方面,有个令人多少感到惊讶的结果就是美式买权的价格和同样的欧式买权的价格刚好相同,即

$$C^A = C^E.$$

我们同时假定股票不分红. 但是,一些考虑揭示了原因. 即拥有欧式买权意味着拥有者能够在任何时间借入标的股票,在 T 时刻用买权来对冲股票的空头,购买需偿还的股票的价格最多为 K. 这就保护了美式买权不能提前执行.

具体地,假设 $C^A > C^E$ 并考虑如下的初始投资组合:

1) -1 份美式买权 @ C^A.

2) 1 份欧式买权 @ C^E.

3) $C^A - C^E > 0$ 份债券 @ 1.

该组合的初始价值为 0. 如上面所提及的, 欧式买权的拥有者能阻止对美式买权的提前执行, 因为我们总能借一股股票来应付美式买权的提前执行.

特别地, 如果买权都没有被执行, 那么有来自无风险头寸的一个保证利润. 如果美式买权在 T 时刻被执行, 那么我们也可相应地执行欧式买权, 同样可以从债券上得到一笔净利润. 最后, 如果美式买权在时刻 $t < T$ 被执行, 那么我们能够借入一股股票来平仓该买权. 这时的组合变为:

1) -1 股 @ S_t.

2) 1 张欧式买权 @ ?.

3) $C^A - C^E > 0$ 份债券 @ e^{rt}.

4) K 美元.

在时刻 T 变为:

1) -1 股 @ S_T.

2) 1 张欧式买权 @ $(K - S_T)^+$.

3) $C^A - C^E > 0$ 份债券 @ e^{rT}.

4) $K\mathrm{e}^{r(T-t)}$ 美元.

现在我们简单地执行买权来填补股票的空头, 或者如果股价已经下跌到执行价 K 以下, 那么直接在市场上买入股票. 用这样方式我们能够使股票的空头得以平仓, 所需花费为 $\min\{K, S_T\}$. 因此最后的利润为

$$(C^A - C^E)\mathrm{e}^{rT} + K\mathrm{e}^{r(T-t)} - \min\{K, S_T\} > 0.$$

只需注意到 $K\mathrm{e}^{r(T-t)} > K$ 即可. 换句话说, 最后执行美式买权比其他任何时候执行都好.

在这方面还需要学习更多有关知识. 也就是说提前执行美式买权是不明智的.

我们可以从几个角度来看. 首先, 如果是想坚持持有股票至到期日 T, 那么在 t 时刻提前执行的话, 在 T 时刻得到相同的投资组合. 与提前执行之间的差别是, 投资者在 t 时刻支付执行价而不是在 T 时刻, 因此在 t 到 T 这段时间里对执行价失去兴趣. 另一方面, 如果是为了想立即卖掉股票, 提前执行期权和卖出股票得到净收入 $S_t - K$(假设为正), 这就是期权的内在价值. 但是期权的市场价值必须至少为该值. 因此, 相对于执行来说, 卖出期权本身更有利.

从另外一个角度来看, 注意到

$$C_A = C^E \geqslant S_0 - K\mathrm{e}^{-rT}.$$

我们希望读者能证明上式后面部分的不等式. 然而, 在 0 时刻提前执行的价值为 S_0-K, 因此只要 $r>0$, 美式买权的价值就大于提前执行所获得价值.

最后, 我们观察到借入股票而不需立即付钱这种选择可使投资者拥有股票. 花费 K 能推迟到最后时间 T, 这时执行期权 (或者在市场上之间买入股票, 如果更便宜的话) 来平仓股票的空头.

定理 4　*假定标的股票不分红, 对于各条款都相同的美式买权和欧式买权来说, 有*

$$C_A=C_E.$$

对于各条款都相同的美式卖权和欧式卖权来说, 有

$$P_A\geqslant P_E.$$

可能是严格不等式. 而且, 提前执行美式买权是不明智的.

练　习　4

如果一个远期合约的标的资产在该合约生命周期内提供 1 美元的收入, 那么合约的多头投资者将损失 1 美元, 而现货持有投资者将得到 1 美元的收入. 这引出了先前的无套利观点.

1. 假设从标的资产获得收入的现值为 I. 在这种情况下的损益是什么? 假设无风险收益的连续复合利率为 r.

2. 在该情形下两种策略的损益是多少?

3. 证明无套利假设意味着有

$$F_{0,T}=(S_0-I)\mathrm{e}^{rT}.$$

我们已经知道在远期合约最简单的情形下 (远期合约没产生收入), 在 0 时刻无套利远期价格为

$$F_{0,T}=S_0\mathrm{e}^{rT}.$$

然而, 我们是在完美市场这个理想的假设下得到该公式的. 我们来看, 如果没有完美市场这个限制会发生什么. 特别地, 假设借贷利率不同, 在现实生活一般总是这样. 记投资者借出利率为 r_l, 借入利率为 r_b. 当然对单个投资者来说有 $r_l<r_b$.

4. 在上面这些条件下, 损益是什么? 假设无风险收益的连续复合利率为 r.

5. 在该情形下两种策略的损益是多少?

6. 证明无套利假设意味着有

$$S_0 e^{r_l T} \leqslant F_{0,T} \leqslant S_0 e^{r_b T}.$$

提示：为了避免套利, 两种策略必须到期收益为非正的损益.

在练习 6 中给出的上下界为无套利边界 (no-arbitrage bounds), 由无套利得出的期货价格的价值范围是无套利价差的 (no-arbitrage spread). 因此, 在不完美市场中, 无套利意味着期货价格可为一定范围内的任一值.

期权价格的上界

7. 利用无套利原理证明：欧式买权或美式买权的初始价格小于标的股票的初始价格, 即

$$C^E \leqslant S_0,$$
$$C^A \leqslant S_0.$$

8. 利用无套利证明：

$$P^E \leqslant K e^{-rT},$$
$$P^A \leqslant K.$$

9. 证明：

$$S_0 - K e^{-rT} - d_0 \leqslant C^E,$$
$$K e^{-rT} + d_0 - S_0 \leqslant P^E.$$

10. a) 证明：对不分红的股票有

$$S_0 - K e^{-rT} \leqslant C^A,$$
$$K - S_0 \leqslant P^A.$$

b) 如果股票支付红利, 其红利的现值为 d_0, 则

$$\max\{S_0 - K e^{-rT} - d_0, S_0 - K\} \leqslant C^A,$$
$$\max\{K e^{-rT} + d_0 - S_0, K - S_0\} \leqslant P^A.$$

第 5 章　概率论 2: 离散概率

在这一章中, 我们讨论有限概率空间上的一些内容, 在下一章的离散时间模型中要用到这部分内容.

5.1　条 件 概 率

当关于一次试验的附加信息有用时, 可用条件概率这个概念将附加信息考虑进去. 想法是通过一种与原概率测度成比例的方式, 将 Ω 上的所有概率 "集中" 到集合 E 上.

定义 1　设 $(\Omega, \mathbb{P})$ 是一个概率空间, 令 E 是概率为 $\mathbb{P}(E) > 0$ 的事件. 那么对任何事件 A, A 关于 E 的条件概率为

$$\mathbb{P}(A|\ E) = \frac{\mathbb{P}(A \cap E)}{\mathbb{P}(E)}.$$

记号 $\mathbb{P}(A|\ E)$ 读作 "A 关于 E 的概率". 注意我们不需要担心 $\mathbb{P}(E) = 0$ 这种情况, 因为没有必要去问关于一个不可能发生的事件的条件概率.

在一个事件上的条件作用允许我们在 Ω 上去定义一个新的 "条件" 概率测度.

定理 1　设 $(\Omega, \mathbb{P})$ 是一个概率空间, 令 E 是概率为 $\mathbb{P}(E) > 0$ 的事件. 那么如下定义的集合函数 $\mathbb{P}_E$:

$$\mathbb{P}_E(A) = \mathbb{P}(A|\ E)$$

是 Ω 上的一个概率测度, 此时有 $\mathbb{P}_E(E) = 1$.

证明　为了证明 $\mathbb{P}_E$ 是 Ω 上的一个概率测度, 我们必须验证一些事实. 首先, $\mathbb{P}$ 的单调性暗示着

$$0 \leqslant \mathbb{P}(A \cap E) \leqslant \mathbb{P}(E).$$

所以 $0 \leqslant \mathbb{P}(A|\ E) \leqslant 1$, 即

$$0 \leqslant \mathbb{P}_E(A) \leqslant 1.$$

也有

$$\mathbb{P}_E(\Omega) = \mathbb{P}(\Omega|\ E) = \frac{\mathbb{P}(\Omega \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E)}{\mathbb{P}(E)} = 1.$$

最后, 如果 $A \cap B = \varnothing$, 那么 $A \cap E$ 和 $B \cap E$ 也是不相交的, 从而有

$$
\begin{aligned}
\mathbb{P}_E(A \cup B) = \mathbb{P}(A \cup B|\ E) &= \frac{\mathbb{P}((A \cup B) \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}((A \cap E) \cup (B \cap E))}{\mathbb{P}(E)} \\
&= \frac{P(A \cap E) + \mathbb{P}(B \cap E)}{\mathbb{P}(E)} = \frac{P(A \cap E)}{\mathbb{P}(E)} + \frac{\mathbb{P}(B \cap E)}{\mathbb{P}(E)} \\
&= \mathbb{P}(A|\ E) + \mathbb{P}(B|\ E) = \mathbb{P}_E(A) + \mathbb{P}_E(B).
\end{aligned}
$$

这就完成了证明.

用条件概率来表示全概率公式会呈现出非常好的形式.

定理 2(全概率公式)　设 Ω 是一个样本空间, $E_1, \cdots, E_n$ 形成 Ω 的一个划分. 假如对所有的 k 有 $\mathbb{P}(E_k) \neq 0$, 那么对任意 $A \in \Omega$, 有

$$
\mathbb{P}(A) = \sum_{k=1}^{n} \mathbb{P}(A|\ E_k)\mathbb{P}(E_k).
$$

5.2　划分和可测性

为了方便, 我们重述划分的定义.

定义 2　设 Ω 是非空集. Ω 的一个划分是由 Ω 中的非空子集形成的群集 $\mathcal{P} = \{B_1, \cdots, B_k\}$, 非空子集称为划分的块, 满足如下性质:

1) 块是两两不相交的

$$
B_i \cap B_j = \varnothing.
$$

2) 所有块的并等于 Ω

$$
B_1 \cup \cdots \cup B_k = \Omega.
$$

图 1 给出了集合 Ω 的一个划分.

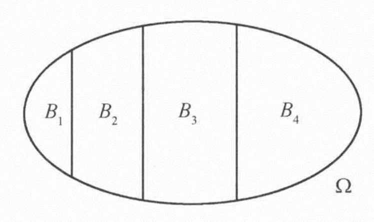


图 1　Ω 的一个划分

我们也使用划分的细分这个概念.

定义 3　设 $\mathcal{P} = \{B_1, \cdots, B_k\}$ 是 Ω 的一个划分. 则称由分解 $\mathcal{P}$ 中某些块而得到的划分 $\mathcal{Q} = \{C_1, \cdots, C_n\}$ 为 $\mathcal{P}$ 的一个细分. 因此, $\mathcal{Q}$ 是 $\mathcal{P}$ 的一个细分, 如果有 $\mathcal{Q}$

的每个块被包含在 $\mathcal{P}$ 中的某些块中, 或等价地有 $\mathcal{P}$ 的每个块是 $\mathcal{Q}$ 中一些块的并. 我们将此关系记为 $\mathcal{P} \prec \mathcal{Q}$.

注意, 当我们说一个集合 A 是群集 $\{B_1, \cdots, B_n\}$ 中的集合的并集时, 包含了 A 是单个集合 B_k 的并这种可能性, 即 $A = B_k$.

图 2 画出了图 1 中那个划分的细分. 注意到

$$\begin{aligned}
B_1 &= C_1 \cup C_2, \\
B_2 &= C_3 \cup C_4 \cup C_5, \\
B_3 &= C_6, \\
B_4 &= C_7.
\end{aligned}$$

所以每个块 B_i 都是块 C_j 的并.

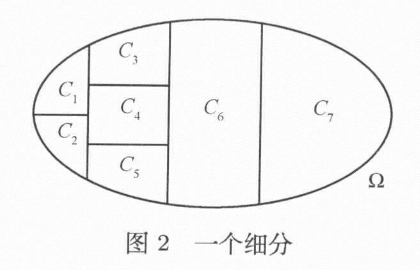


图 2　一个细分

由随机变量定义的划分

假设 X 是一个随机变量, 不像普通函数那样, 通常不用 $X^{-1}(B)$, 而是用

$$\{X \in B\}$$

来记集合 B 在 X 下的原象. 同样, 通常用下面的表达式

$$\{X = x\}$$

来代替 $X^{-1}(x)$. 我们也提醒读者, X 的取值集合 $\{x_1, \cdots, x_n\}$ 叫做 X 的象, 记为 $\mathrm{im}(X)$.

有限样本空间 Ω 上的任意随机变量 X 定义了 Ω 上的一个划分 $\mathcal{P}_X$, 如图 3 所示.

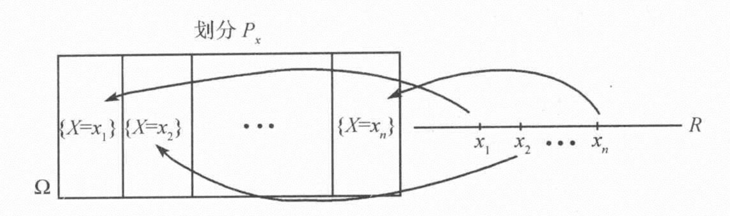


图 3　由随机变量定义的划分

定义 4　设 X 是 Ω 上的一个随机变量, 且有

$$\operatorname{im}(X)=\{x_1,\cdots,x_n\}.$$

那么 X 定义了 Ω 的一个划分, 它的块是 $\operatorname{im}(X)$ 中元素的原象, 即

$$\mathcal{P}_X=\{\{X=x\}|\ x\in\operatorname{im}(X)\}=\{\{X=x_1\},\cdots,\{X=x_n\}\},$$

$\mathcal{P}_X$ 就称为 X 定义的划分.

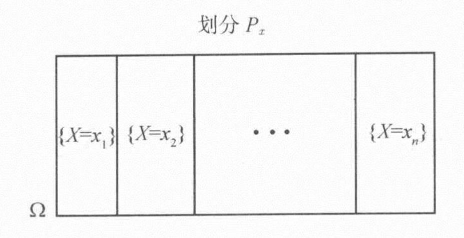
随机变量关于划分的可测性

由随机变量定义的划分 $\mathcal{P}_X$ 有一个非常重要的性质: X 在 $\mathcal{P}_X$ 的块 $\{X=x\}$ 上是常数. 事实上, X 在 $\{X=x\}$ 上取常值 x. 我们用 X 是 $\mathcal{P}_X$ 可测的来表达这种性质.

定义 5　设 $\mathcal{P}$ 是 Ω 的任一划分. Ω 上的随机变量 X 是 $\mathcal{P}_X$ 可测的, 如果 X 在 $\mathcal{P}$ 的每个块上是常数.

X 关于 $\mathcal{P}_X$ 还有其他一个很明显的性质, 那就是 X 不仅在 $\mathcal{P}_X$ 的每个块上是常数, 而且这些常数全不相同.

现在, 给定一个非常数的随机变量, 存在许多划分 $\mathcal{Q}$, 使得 X 在 $\mathcal{Q}$ 的每个块上都是常数, 即 X 是 $\mathcal{Q}$ 可测的. 但是, $\mathcal{P}_X$ 是唯一使得 X 在每个块上取不同的常数值的划分. 我们可以非常简单地刻画使 X 可测的所有划分 (见图 4).



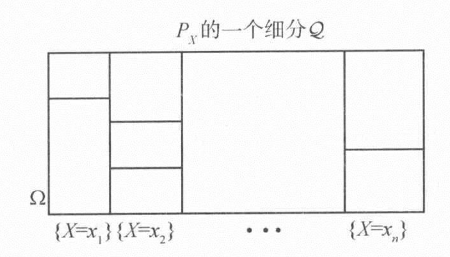


图 4　P_x 的细分

定理 3 设 X 是 Ω 上的一个随机变量.

1) 那么 X 是 $\mathcal{Q}$ 可测的当且仅当 $\mathcal{Q}$ 是 $\mathcal{P}_X$ 的一个细分.

2) $\mathcal{P}_X$ 是使得 X 可测的最粗糙划分, 是使得 X 可测且在每个块上取不同的值的唯一划分.

证明 关于 1). 如果 X 是 $\mathcal{Q}$ 可测的且 $\mathcal{Q} = \{B_1, \cdots, B_k\}$, 那么对任意 $\omega \in B_i$, 有

$$B_i \subseteq \{X = X(\omega)\}.$$

所以 $\mathcal{Q}$ 是 $\mathcal{P}_X$ 的一个细分. 反向很显然.

2) 的证明留给读者.

下面这个非常重要的定理指出了随机变量 X 和另外一个 $\mathcal{P}_X$ 可测的随机变量 Y 之间非常强的联系.

定理 4 设 X 和 Y 是两个随机变量. 那么 Y 是 $\mathcal{P}_X$ 可测的当且仅当 Y 是 X 的函数, 即当且仅当存在一个函数 $f: \mathbb{R} \to \mathbb{R}$, 使得

$$Y = f(X).$$

证明 我们知道 Y 在划分 $\mathcal{P}_X = \{B_1, \cdots, B_k\}$ 的每个块上是常数. 不妨假定

$$Y(B_i) = \{y_i\}.$$

当然, X 在 $\mathcal{P}_X$ 的每个块上也是常数, 故令

$$X(B_i) = \{x_i\}.$$

定义 f 如下

$$f(x_i) = y_i.$$

从而对 $\omega \in B_i$, 有

$$f(X)(\omega) = f(X(\omega)) = f(x_i) = y_i = Y(\omega).$$

因此 $Y = f(X)$, 必要性得证. 充分性的证明容易很多, 我们将它留作练习.

划分与独立

我们另外简单地看独立性这个概念. 下面重新给出它的定义.

定义 6 称 $(\Omega, \mathbb{P})$ 中的事件 E 和 F 独立, 如果

$$\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F).$$

称事件 $E_1, \cdots, E_k$ 独立, 如果对这些事件的任何子事件类 $E_{i_1}, \cdots, E_{i_m}$ 有

$$\mathbb{P}(E_{i_1} \cap \cdots \cap E_{i_m}) = \mathbb{P}(E_{i_1}) \cdots \mathbb{P}(E_{i_m}).$$

可以将独立性的定义推广到一族事件类上.

定义 7 称事件类 $\mathcal{C}_1, \cdots, \mathcal{C}_k$ 独立, 如果对任何选择 $\mathbb{E}_i \in \mathcal{C}_i$ 有事件 $E_1, \cdots, E_k$ 独立.

当事件类是 Ω 的划分时, 我们有理由应用该定义. 事实上, 我们回顾一下, 称随机变量 $X_1, \cdots, X_n$ 是独立的, 如果对所有的 $x_1, \cdots, x_n \in \mathbb{R}$, 有

$$\mathbb{P}(X_1 = x_1, \cdots, X_n = x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i).$$

该定义也可用划分的形式重新表示如下.

定理 5 随机变量 $X_1, \cdots, X_n$ 是独立的当且仅当划分 $\mathcal{P}(X_1), \cdots, \mathcal{P}(X_n)$ 是独立的事件类.

证明 从 $\mathcal{P}(X_i)$ 的块刚好是集合 $\{X_i = x_i\}$ 这个事实出发, 定理很快就得证.

5.3 代 数

我们已经知道在有限样本空间中, 划分与随机变量密切联系在一起. 碰巧的是, 划分的概念不能推广到非有限样本空间上. 正因为如此, 我们需要另外一个概念 —— 代数.

我们不会直接用到该节讨论的内容, 因为对离散事件模型的分析来说划分已经足够了, 而且划分比代数更直观. 但是, 我们想在这里讨论代数, 是因为它们在划分和 σ 代数之间起到了桥梁作用. 在连续时间定价模型和布莱克 – 舒尔斯期权定价公式的分析中需要用到 σ 代数.

定义 8 设 Ω 是一个非空集. 称由 Ω 的子集形成的事件类 $\mathcal{A}$ 是一个代数, 如果它满足下面的性质:

1) (空集在$\mathcal{A}$中)

$$\varnothing \in \mathcal{A}.$$

2) ($\mathcal{A}$对补集运算封闭)

$$A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}.$$

3) ($\mathcal{A}$对并集运算封闭)

$$A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}.$$

不难证明, 任何代数在交集和差集运算下也是封闭的, 即

$$A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A},$$

$$A, B \in \mathcal{A} \Rightarrow A \backslash B \in \mathcal{A}.$$

下面的概念非常有用. 它明确了一个代数 $\mathcal{A}$ 中"最小的"非空集合的概念.

定义 9　设 $\mathcal{A}$ 是 Ω 上的一个代数. $\mathcal{A}$ 的一个原子就是一个非空集合 $A \in \mathcal{A}$, 且具有性质: 在 $\mathcal{A}$ 中不存在 A 的非空真子集.

划分与代数

从 Ω 的一个划分 $\mathcal{P}$ 出发, 通过简单地取 $\mathcal{P}$ 中元素的所有可能的并, 我们可以生成一个代数 $\mathcal{A}(\mathcal{P})$. 反之也是可能的: 从有限样本空间的一个代数出发, 我们能得到一个划分.

定理 6　设 Ω 是非空有限集.

1) 对 Ω 的任意划分 $\mathcal{P}$, 集合

$$\mathcal{A}(\mathcal{P}) = \{C \subseteq \Omega | \ C = \varnothing \text{或者} C = \mathcal{P} \text{中块的并}\}$$

是一个代数, 称为由 $\mathcal{P}$ 生成的代数.

2) 如果 $\mathcal{A}$ 是 Ω 上的一个代数, 那么由 $\mathcal{A}$ 的所有原子构成的集合

$$\mathcal{P}(\mathcal{A}) = \{\mathcal{A} \text{的所有原子}\}$$

是 Ω 的一个划分, 称为由 $\mathcal{A}$ 定义的划分.

证明　我们只证明 2), 将 1) 留作练习. 令

$$\mathcal{P} = \{A_1, \cdots, A_k\}$$

由 $\mathcal{A}$ 中所有不同的原子组成. 我们必须证明 $\mathcal{P}$ 是 Ω 的一个划分. 根据原子的定义知 $\mathcal{P}$ 非空. 如果 $A_i \cap A_j \neq \varnothing$ $(i \neq j)$, 那么 $A_i \cap A_j$ 是 $\mathcal{A}$ 中的非空元素, 且是 A_i 的真子集, 这是不可能的. 因此, 原子是两两不相交的. 最后, 假设 $\omega \in \Omega$. $\mathcal{A}$ 中所有包含 ω 的元素的交 I 也是 $\mathcal{A}$ 中包含 ω 的一个元素. 而且, I 非空和不存在 $\mathcal{A}$ 中的真子集, 因为这样的真子集会成为定义 I 的交中的一部分. 因此, I 实际上是 $\mathcal{A}$ 的原子. 这表明 Ω 的每个元素都包含在 $\mathcal{A}$ 的某个原子里. 所以 $\mathcal{A}$ 的所有原子的并等于 Ω. 从而 $\mathcal{P}$ 是 Ω 的一个划分.

我们目前讨论的主题是, 在有限样本空间中划分和代数两个概念在某种意义上是等价的, 即

$$\Omega \text{的划分} \Leftrightarrow \Omega \text{上的代数}.$$

意思是说, 虽然这些概念确实不相同, 但是对于代数来说有类似划分的所有陈述, 反之一样. 换个角度说, 不管在划分的内容上建立了什么样的理论, 我们也能在代数的内容上建立同样的理论, 反之一样.

两个概念的准确联系由定理 6 中描述的对应给出

$$\mathcal{Q} \to \mathcal{A}(\mathcal{Q}) \quad [\text{划分 - 划分生成的代数}],$$
$$\mathcal{B} \to \mathcal{P}(\mathcal{B}) \quad [\text{代数 - 由代数定义的划分}].$$

第一个对应说明取 Ω 的任何划分可生成一个代数, 第二个对应说明取任何代数可产生一个划分. 事实上, 一个对应是另一个对应的逆 (所以是一一对应).

为了看清这一点, 假设 $\mathcal{Q}$ 是 Ω 的一个划分, $\mathcal{A}(\mathcal{Q})$ 是 $\mathcal{Q}$ 生成的代数. 因此, $\mathcal{Q}$ 的块刚好是 $\mathcal{A}(\mathcal{Q})$ 的原子, 即

$$\mathcal{P}(\mathcal{A}(\mathcal{Q})) = \mathcal{Q}.$$

同样地, 如果我们从 Ω 的一个代数 $\mathcal{B}$ 出发, 那么划分 $\mathcal{P}(\mathcal{B})$ 由 $\mathcal{B}$ 的原子组成. 但是 $\mathcal{B}$ 的每个元素都是 $\mathcal{B}$ 的原子的并, 所以 $\mathcal{A}(\mathcal{P}(\mathcal{B}))$ 与 $\mathcal{B}$ 是一样的, 即

$$\mathcal{A}(\mathcal{P}(\mathcal{B})) = \mathcal{B}.$$

这就表明两个对应是一一对应的, 互为逆. 这是两个概念间非常紧密的联系.

下一定理加强了划分与代数的关系. 道出了划分的细分概念对应着代数包含的概念. 注意在定理中有两种陈述. 这两种陈述表达了相同的事情：一种是从划分的观点看, 一种是从代数的观点看. 记得我们用 $\mathcal{P} \prec \mathcal{Q}$ 表示 $\mathcal{Q}$ 是 $\mathcal{P}$ 的一个细分.

定理 7 *1) 设 $\mathcal{P}$ 和 $\mathcal{Q}$ 是 Ω 的划分, 那么*

$$\mathcal{A}(\mathcal{P}) \subseteq \mathcal{A}(\mathcal{Q}) \Leftrightarrow \mathcal{P} \prec \mathcal{Q}.$$

2) 设 $\mathcal{A}$ 和 $\mathcal{B}$ 是 Ω 上的代数, 那么

$$\mathcal{A} \subseteq \mathcal{B} \Leftrightarrow \mathcal{P}(\mathcal{A}) \prec \mathcal{P}(\mathcal{B}).$$

证明 我们只需证明 1)(为什么?). 假设 $\mathcal{A}(\mathcal{P}) \subseteq \mathcal{A}(\mathcal{Q})$, 因而 $\mathcal{P}$ 的块是 $\mathcal{A}(\mathcal{P})$ 中的原子. 如果 A 是 $\mathcal{P}$ 的一个原子, 那么它也在 $\mathcal{A}(\mathcal{Q})$ 中. 从而它是 $\mathcal{Q}$ 中块的并. 换句话说, $\mathcal{P}$ 的每个块是 $\mathcal{Q}$ 中块的并, 所以 $\mathcal{Q}$ 是 $\mathcal{P}$ 的一个细分.

另一方面, 假设 $\mathcal{Q}$ 是 $\mathcal{P}$ 的一个细分. 那么 $\mathcal{P}$ 的任何块都是 $\mathcal{Q}$ 中块的并, 从而 $\mathcal{A}(\mathcal{P})$ 中任意元素 $\mathcal{P}$ 的块的并也是 $\mathcal{Q}$ 中块的并, 因此属于 $\mathcal{A}(\mathcal{Q})$, 得到 $\mathcal{A}(\mathcal{P}) \subseteq \mathcal{A}(\mathcal{Q})$.

随机变量生成的代数

我们已经看到了 Ω 的划分和 Ω 上的代数之间的紧密关系. 现在是把随机变量引进来的时候了.

正如一个随机变量 X 定义了 Ω 上的一个划分 $\mathcal{P}_X$,

$$\mathcal{P}_X = \{\{X = x\} |\ x \in \operatorname{im}(X)\},$$

由 $\mathrm{im}(X)$ 中元素的原象构成. 随机变量也定义了 Ω 上的一个代数 $\mathcal{A}_X$, 由 $\mathrm{im}(X)$ 的子集的原象构成.

定义 10　设 X 是 Ω 上的一个随机变量. 那么 X 定义了 Ω 上的一个代数, 其元素为 $\mathrm{im}(X)$ 子集的原象, 即

$$\mathcal{A}_X = \{\{X \in B\} |\ B \subseteq \mathrm{im}(X)\}.$$

容易看到 $\mathcal{P}_X$ 和 $\mathcal{A}_X$ 是相联的 (见图 5).

实际上, $\mathcal{A}_X$ 就是 $\mathcal{P}_X$ 生成的代数, 即

$$\mathcal{A}_X = \mathcal{A}(\mathcal{P}_X).$$

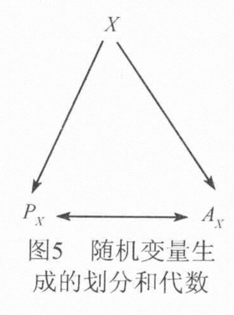


图5　随机变量生成的划分和代数

为了看清这点, 注意到如果 $B = \{b_1, \cdots, b_m\}$ 是 $\mathrm{im}(X)$ 的一个子集, 那么

$$\{X \in B\} = \bigcup_{i=1}^{m} \{X = b_i\} \in \mathcal{A}(\mathcal{P}_X).$$

所以 $\mathcal{A}_X \subseteq \mathcal{A}(\mathcal{P}_X)$. 但是 $\mathcal{A}(\mathcal{P}_X)$ 是包含 $\mathcal{P}_X$ 所有块 $\{X = b_i\}$ 的最小代数, 因此 $\mathcal{A}_X = \mathcal{A}(\mathcal{P}_X)$.

定理 8　设 X 是有限样本空间 Ω 上的随机变量. 那么 X 生成的代数就是 $\mathcal{P}_X$ 生成的代数, 即

$$\mathcal{A}_X = \mathcal{A}(\mathcal{P}_X),$$

并且 X 定义的划分就是 $\mathcal{A}_X$ 定义的划分, 即

$$\mathcal{P}_X = \mathcal{P}(\mathcal{A}_X).$$

随机变量关于代数的可测性

我们已经定义了随机变量 X 关于 Ω 的划分 $\mathcal{P}$ 的可测性. 意思是 X 在 $\mathcal{P}$ 的块上为常数. 也看到 X 是 $\mathcal{P}$ 可测的当且仅当 $\mathcal{P}$ 是 $\mathcal{P}_X$ 的一个细分. 现在我们转到 X 关于 Ω 上一个代数的可测性上来.

不应该感到奇怪, 我们想要该定义使得随机变量 X 是 $\mathcal{A}(\mathcal{P})$ 可测的当且仅当它是 $\mathcal{P}$ 可测的 (即在$\mathcal{P}$的每个块上是常数). 事实上, 因为任何代数 $\mathcal{A}$ 可由一个划分生成, 即

$$\mathcal{A} = \mathcal{A}(\mathcal{P}(\mathcal{A})).$$

所以实际上可以用它作为关于代数可测性的定义性质. 该定义如下：设 $\mathcal{A}$ 是 Ω 上的代数, 那么 X 是 $\mathcal{A}$ 可测的, 如果 X 在 $\mathcal{A}$ 的所有原子上都是常量.

虽然上述定义很直观, 但是不标准, 采用它使读者感到不亲切. 为了理解通常的定义, 注意到下面条件等价:

1) X 是 $\mathcal{P}$ 可测的.

2) X 在 $\mathcal{P}$ 的块上为常量.

3) X 在 $\mathcal{A}(\mathcal{P})$ 的原子上为常量.

4) 每个集合 $\{X = x\}$ 都是 $\mathcal{A}(\mathcal{P})$ 原子的并.

5) 每个集合 $\{X = x\}$ 属于 $\mathcal{A}(\mathcal{P})$.

6) 对任何子集 $B \subseteq \mathrm{im}(X)$, 有集合 $\{X \in B\}$ 属于 $\mathcal{A}(\mathcal{P})$.

现在我们已经为标准定义准备好了. 它等价于先前直觉上的定义.

定义 11 设 X 是有限样本空间 Ω 上的随机变量, $\mathcal{A}$ 是 Ω 上的任意代数. 称 X 是 $\mathcal{A}$ 可测的, 如果

$$\{X \in B\} \in \mathcal{A}, \text{对所有的} B \subseteq \mathrm{im}(X)$$

我们已经指出了 X 是 $\mathcal{P}$ 可测的当且仅当 X 是 $\mathcal{A}(\mathcal{P})$ 可测的.

定理 9 设 X 是 Ω 上的随机变量.

1) 如果 $\mathcal{A}$ 是 Ω 上的代数, 那么 X 是 $\mathcal{A}$ 可测的当且仅当它是 $\mathcal{P}(\mathcal{A})$ 可测的.

2) 如果 $\mathcal{P}$ 是 Ω 的一个划分, 那么 X 是 $\mathcal{P}$ 可测的当且仅当它是 $\mathcal{A}(\mathcal{P})$ 可测的.

如刻画随机变量关于划分的可测性那样 ——X 是 $\mathcal{P}$ 可测的当且仅当 $\mathcal{P}_X \prec \mathcal{P}$, 我们也能刻画随机变量关于代数的可测性.

定理 10 Ω 上的随机变量 X 关于代数 $\mathcal{A}$ 是可测的当且仅当 $\mathcal{A}_X \subseteq \mathcal{A}$.

证明 下面的等价陈述可证明该定理.

1) X 是 $\mathcal{A}$ 可测的.

2) X 是 $\mathcal{P}(\mathcal{A})$ 可测的.

3) $\mathcal{P}(\mathcal{A})$ 是 $\mathcal{P}_X$ 的一个细分.

4) $\mathcal{A}(\mathcal{P}(\mathcal{A}))$ 包含 $\mathcal{A}(\mathcal{P}_X)$.

5) $\mathcal{A}$ 包含 $\mathcal{A}_X$.

5.4 条件期望

我们可以结合条件概率和期望来得到条件期望, 条件期望在衍生品定价模型中起着重要作用.

关于一个事件的条件期望

关于一个正概率事件 A 的条件期望很简单 —— 我们采用通常的期望, 只是在

条件概率测度 $\mathbb{P}_A$ 下, 其中 $\mathbb{P}_A$ 的定义为

$$\mathbb{P}_A(B) = \mathbb{P}(B|\ A).$$

定义 12 设 $(\Omega, \mathbb{P})$ 是有限概率空间, 令 A 为概率大于零的事件. 随机变量 X 关于事件 A 的条件期望为

$$\mathcal{E}_{\mathbb{P}}(X|\ A) = \mathcal{E}_{\mathbb{P}_A}(X).$$

记号 $\mathcal{E}_{\mathbb{P}}(X|\ A)$ 读作"X 关于 A 的期望".

运用一点代数学知识就能得到用非条件期望的形式来表示条件期望的有用表达式.

定理 11 设 $(\Omega, \mathbb{P})$ 是有限概率空间, 令 A 为概率大于零的事件. 随机变量 X 关于事件 A 的条件期望为

$$\mathcal{E}_{\mathbb{P}}(X|\ A) = \frac{\mathcal{E}_{\mathbb{P}}(X1_A)}{\mathbb{P}(A)}.$$

其中 1_A 是 A 的示性函数.

证明 我们有

$$\begin{aligned}
\mathcal{E}_{\mathbb{P}_A}(X) &= \sum_{i=1}^{n} X(\omega_i)\mathbb{P}(\omega_i|\ A) \\
&= \sum_{i=1}^{n} X(\omega_i)\frac{\mathbb{P}(\omega_i \cap A)}{\mathbb{P}(A)} \\
&= \frac{1}{\mathbb{P}(A)}\sum_{i=1}^{n} X(\omega_i)\mathbb{P}(\omega_i \cap A) \\
&= \frac{1}{\mathbb{P}(A)}\sum_{i=1}^{n} X(\omega_i)1_A(\omega_i)\mathbb{P}(\omega_i) \\
&= \frac{1}{\mathbb{P}(A)}\mathcal{E}_{\mathbb{P}}(X1_A).
\end{aligned}$$

得证.

上述定理的一个简单推论就是下面的有用结果.

定理 12 如果两个事件 A 和 B 满足 $\mathbb{P}(A \cap B) > 0$, 那么

$$\mathcal{E}_{\mathbb{P}_A}(X|\ B) = \mathcal{E}(X|\ A \cap B).$$

证明　根据前面的定理, 有

$$\begin{aligned}\mathcal{E}_{\mathbb{P}_A}(X|\ B) &= \frac{\mathcal{E}_{\mathbb{P}_A}(X1_B)}{\mathbb{P}_A(B)}\\ &= \frac{\mathcal{E}_{\mathbb{P}}(X1_B1_A)}{\mathbb{P}(A)\mathbb{P}_A(B)}\\ &= \frac{\mathcal{E}_{\mathbb{P}}(X1_{A\cap B})}{\mathbb{P}(A\cap B)}\\ &= \mathcal{E}_{\mathbb{P}}(X|\ A\cap B).\end{aligned}$$

得证.

下面我们在期望方面有个类似全概率公式的定理.

定理 13　设 $\mathcal{P}=\{B_1,\cdots,B_n\}$ 是 Ω 的一个划分. 那么对 Ω 上的任何随机变量 X, 有

$$\mathcal{E}(X)=\sum_{i=1}^{n}\mathcal{E}(X|\ B_i)\mathbb{P}(B_i).$$

而且, 如果 A 是一个概率大于零的事件, 那么

$$\mathcal{E}(X|\ A)=\sum_{i=1}^{n}\mathcal{E}(X|\ B_i\cap A)\mathbb{P}(B_i|\ A).$$

假如我们认为不能定义条件期望的项等于零, 则上面式子中的和有意义. 换个方式说, 如果一个项中的任一因子为零, 那么认为该项也为零, 即

$$\text{没有定义}\cdot 0=0.$$

证明　关于第一部分, 假设

a) $\mathbb{P}(B_i)>0, i=1,\cdots,m.$

b) $\mathbb{P}(B_i)=0, i=m+1,\cdots,n.$

因为

$$X=\sum_{i=1}^{m}X1_{B_i}+\sum_{i=m+1}^{n}X1_{B_i}.$$

取期望得

$$\begin{aligned}\mathcal{E}(X) &= \sum_{i=1}^{m}\mathcal{E}(X1_{B_i})+\sum_{i=m+1}^{n}\mathcal{E}(X1_{B_i})\\ &= \sum_{i=1}^{m}\frac{\mathcal{E}(X1_{B_i})}{\mathbb{P}(B_i)}\mathbb{P}(B_i)\\ &= \sum_{i=1}^{m}\mathcal{E}(X|\ B_i)\mathbb{P}(B_i).\end{aligned}$$

这就证明了第一部分.

对于第二部分, 我们将第一部分的结论应用到条件概率测度 $\mathbb{P}_A$ 上, 有

$$\mathcal{E}_{\mathbb{P}}(X|\ A) = \mathcal{E}_{\mathbb{P}_A}(X) = \sum_{i=1}^{n} \mathcal{E}_{\mathbb{P}_A}(X|\ B_i)\mathbb{P}_A(B_i),$$

其中无定义的项为 0. 除了那些没有定义的项, 即有 $\mathbb{P}_A(B_i) = 0$ 或最终 $\mathbb{P}(A \cap B_i) = 0$(注意$\mathbb{P}_A(B_i)$总是有定义的, 因为假设$A$有正的概率), 其他所有项可以写成

$$\mathcal{E}_{\mathbb{P}_A}(X|\ B_i) = \mathcal{E}_{\mathbb{P}}(X|\ B_i \cap A).$$

即可得证.

关于划分的条件期望

接下来定义随机变量关于样本空间一个划分的条件期望. 不像关于事件的条件期望是一个实数, 关于划分的条件期望是一个随机变量.

我们简要回顾随机变量的通常期望. 当然, 随机变量 X 的期望 $\mathcal{E}(X)$ 是常数. 实际上, 它代表了 X 的最佳可能常数逼近. 用来衡量逼近效果的标准是均值平方误差 (MSE), 其定义为

$$\text{MSE} = \mathcal{E}[(X - c)^2],$$

其中 c 是常数. 碰巧的是, 在所有常数 c 中, 均值平方误差最小当且仅当 c 是期望值 μ_X, 即

$$\mathcal{E}[(X - \mu_X)^2] \leqslant \mathcal{E}[(X - c)^2],$$

取等号当且仅当 $c = \mu_X$. 为了证明这点, 我们写成

$$\begin{aligned}\mathcal{E}[(X - c)^2] &= \mathcal{E}[\{(X - \mu_X) + (\mu_X - c)\}^2] \\ &= \mathcal{E}[(X - \mu_X)^2] + \mathcal{E}[(X - \mu_X)(\mu_X - c)] + \mathcal{E}[(\mu_X - c)^2].\end{aligned}$$

但是中间项为 0(为什么?), 所以

$$\mathcal{E}[(X - c)^2] = \mathcal{E}[(X - \mu_X)^2] + \mathcal{E}[(\mu_X - c)^2] \geqslant \mathcal{E}[(X - \mu_X)^2],$$

取 “=” 当且仅当 $c = \mu_X$.

现在, 我们希望关于划分 $\mathbb{P}$ 的期望 (在划分的每个块上为常量) 是 X 的最佳逼近. 要点是, 如果给定划分的一个块 B, 那么对 X 来说, 我们在 B 上能得到比通常期望值更好的常数逼近. 事实上, 我们用条件期望 $\mathcal{E}(X|\ B)$ 来得到最佳常数逼近. 该想法如图 6 所示, 图中的随机变量 X 定义在区间 $[a, b]$ 上 (我们选择这个例子是因为描绘出这样一个区间上的条件期望比有限样本空间容易).

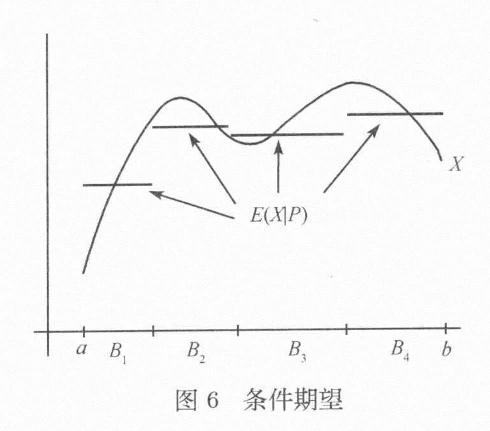


图 6　条件期望

现在 $\mathcal{E}_{\mathbb{P}}(X|\ \mathcal{P})$ 的定义应该很清楚了：它在 $\mathcal{P}$ 的每个块 B_i 上的值为 $\mathcal{E}(X|\ B_i)$. 因此, 我们可以定义 $\mathcal{E}_{\mathbb{P}}(X|\ \mathcal{P})$ 为 $\mathcal{P}$ 的每个块的示性函数的线性组合.

定义 13　设 $(\Omega, \mathbb{P})$ 是一个有限概率空间, 令 $\mathbb{P} = \{B_1, \cdots, B_n\}$ 是 Ω 的一个划分且对所有 i, 有 $\mathbb{P}(B_i) > 0$. 随机变量 X 关于 $\mathbb{P}$ 的条件期望是一个随机变量

$$\mathcal{E}_{\mathbb{P}}(X|\ \mathcal{P}) : \Omega \to \mathbb{R}.$$

其定义为

$$\mathcal{E}_{\mathbb{P}}(X|\ \mathcal{P}) = \mathcal{E}_{\mathbb{P}}(X|\ B_1)1_{B_1} + \cdots + \mathcal{E}_{\mathbb{P}}(X|\ B_n)1_{B_n}.$$

特别地, 对任意 $\omega \in \Omega$, 有

$$\mathcal{E}_{\mathbb{P}}(X|\ \mathcal{P})(\omega) = \mathcal{E}_{\mathbb{P}}(X|\ [\omega]_{\mathcal{P}}),$$

其中 $[\omega]_{\mathcal{P}}$ 表示 $\mathcal{P}$ 中包含 ω 的块.

这里有条件期望在逼近 X 方面的正式陈述.

定理 14　在 $\mathcal{P}$ 的每个块上为常数的所有函数中, 即在所有 $\mathcal{P}$ 可测的随机变量中, 在均值平方误差意义下, 随机变量 $\mathcal{E}(X|\ \mathcal{P})$ 是 X 的最佳逼近. 其意思是对所有 $\mathcal{P}$ 可测的随机变量 Y, 有

$$\mathcal{E}[(X - \mathcal{E}(X|\ \mathcal{P}))^2] \leqslant \mathcal{E}[(X - Y)^2],$$

取 "=" 当且仅当 $Y = \mathcal{E}(X|\ \mathcal{P})$.

证明　为了可读性, 记

$$\mathcal{E}(X|\ \mathcal{P}) = \mu_{X|\ \mathcal{P}}.$$

从 $\mathcal{P}$ 可测的任意随机变量 Y 出发,

$$\begin{aligned}\mathcal{E}[(X-Y)^2]=&\mathcal{E}[\{(X-\mu_{X|\mathcal{P}})+(\mu_{X|\mathcal{P}}-Y)\}^2]\\=&\mathcal{E}[(X-\mu_{X|\mathcal{P}})^2]+\mathcal{E}[(\mu_{X|\mathcal{P}}-Y)^2]\\&+\mathcal{E}[(X-\mu_{X|\mathcal{P}})(\mu_{X|\mathcal{P}}-Y)].\end{aligned}$$

现在证明最后一项为 0. 这可利用定理 13 做到. 假设 $\mathbb{P}=\{B_1,\cdots,B_n\}$, 有

$$\mathcal{E}[(X-\mu_{X|\mathcal{P}})(\mu_{X|\mathcal{P}}-Y)]=\sum_{i=1}^{n}\mathcal{E}[(X-\mu_{X|\mathcal{P}})(\mu_{X|\mathcal{P}}-Y)|\ B_i]\mathbb{P}(B_i).$$

现将注意力集中到表达式

$$\mathcal{E}[(X-\mu_{X|\mathcal{P}})(\mu_{X|\mathcal{P}}-Y)|\ B_i]=\frac{1}{\mathbb{P}(B_i)}\mathcal{E}[(X-\mu_{X|\mathcal{P}})(\mu_{X|\mathcal{P}}-Y)1_{B_i}],$$

该项有 $\mathbb{P}(B_i)>0$(其他项等于 0). 因为 Y 是 $\mathcal{P}$ 可测的, 所以随机变量 $\mu_{X|\mathcal{P}}-Y$ 在 B_i 上为常量, 从而可将它从期望算子里移出来得到

$$\frac{1}{\mathbb{P}(B_i)}(\mu_{X|\mathcal{P}}-Y)\mathcal{E}[(X-\mu_{X|\mathcal{P}})1_{B_i}].$$

又因为 $\mu_{X|\mathcal{P}}1_{B_i}=\mathcal{E}(X|\ B_i)1_{B_i}$, 故有

$$\begin{aligned}\frac{1}{\mathbb{P}(B_i)}\mathcal{E}[(X-\mu_{X|\mathcal{P}})1_{B_i}]=&\frac{1}{\mathbb{P}(B_i)}[\mathcal{E}(X1_{B_i})-\mathcal{E}(\mu_{X|\mathcal{P}}1_{B_i})]\\=&\frac{1}{\mathbb{P}(B_i)}[\mathcal{E}(X1_{B_i})-\mathcal{E}(\mathcal{E}(X|\ B_i)1_{B_i})]\\=&\mathcal{E}(X|\ B_i)-\mathcal{E}(X|\ B_i)\\=&0.\end{aligned}$$

因此, 我们证明了最后一项为 0, 从而

$$\mathcal{E}[(X-Y)^2]=\mathcal{E}[(X-\mu_{X|\mathcal{P}})^2]+\mathcal{E}[(\mu_{X|\mathcal{P}}-Y)^2]\geqslant\mathcal{E}[(X-\mu_{X|\mathcal{P}})^2],$$

取 “=” 当且仅当 $Y=\mathcal{E}(X|\ \mathcal{P})$. 定理得证.

下面的定理给出了条件期望的一些重要性质.

定理 15 设 $(\Omega,\mathbb{P})$ 是一个有限概率空间, 令 $\mathbb{P}=\{B_1,\cdots,B_n\}$ 是 Ω 的一个划分且对所有 i, 有 $\mathbb{P}(B_i)>0$. 条件期望 $\mathcal{E}_{\mathbb{P}}(X|\ \mathcal{P})$ 有下列性质:

1) 函数 $\mathcal{E}(\cdot|\ \mathcal{P})$ 是线性的, 也就是说, 对随机变量 X 和 Y 及实数 a 和 b, 有

$$\mathcal{E}(aX+bY|\ \mathcal{P})=a\mathcal{E}(X|\ \mathcal{P})+b\mathcal{E}(Y|\ \mathcal{P}).$$

2) 条件期望满足

$$\mathcal{E}(\mathcal{E}(X|\ \mathcal{P})) = \mathcal{E}(X).$$

3) 可用唯一的随机变量 Y 来刻画条件期望 $\mathcal{E}(X|\ \mathcal{P})$, 其中 Y 是 $\mathcal{P}$ 可测的, 且满足

$$\mathcal{E}(Y1_{B_i}) = \mathcal{E}(X1_{B_i}),$$

对 $\mathcal{P}$ 的所有块 B_i.

4) 如果 Y 是 $\mathcal{P}$ 可测的随机变量, 那么

$$\mathcal{E}(YX|\ \mathcal{P}) = Y\mathcal{E}(X|\ \mathcal{P}).$$

5) 如果 X 是 $\mathcal{P}$ 可测的, 那么

$$\mathcal{E}(X|\ \mathcal{P}) = X.$$

6) (平滑性) 如果 $\mathcal{Q}$ 是一个比 $\mathcal{P}$ 更细的划分, 那么有

$$\mathcal{E}(\mathcal{E}(X|\ \mathcal{P})|\ \mathcal{Q}) = \mathcal{E}(X|\ \mathcal{P}) = \mathcal{E}(\mathcal{E}(X|\ \mathcal{Q})|\ \mathcal{P}).$$

意思是说, 如果我们关于 $\mathcal{P}$ 和 $\mathcal{Q}$ 取期望而不管次序如何, 那么唯一起作用的就是关于较粗糙那个划分的期望.

7) 如果 X 和 $\mathcal{P}$ 是独立的, 即 $\mathcal{P}_X$ 和 $\mathcal{P}$ 是独立的, 那么

$$\mathcal{E}(X|\ \mathcal{P}) = \mathcal{E}(X).$$

证明　1) 我们有

$$\begin{aligned}
\mathcal{E}(aX + bY|\ \mathcal{P}) &= \sum_{i=1}^{k} \mathcal{E}(aX + bY|\ B_i)1_{B_i} \\
&= \sum_{i=1}^{k} \frac{\mathcal{E}((aX + bY)1_{B_i})}{\mathbb{P}(B_i)} 1_{B_i} \\
&= \sum_{i=1}^{k} \frac{\mathcal{E}(aX1_{B_i} + bY1_{B_i})}{\mathbb{P}(B_i)} 1_{B_i} \\
&= \sum_{i=1}^{k} \frac{a\mathcal{E}(X1_{B_i}) + b\mathcal{E}(Y1_{B_i})}{\mathbb{P}(B_i)} 1_{B_i} \\
&= a\sum_{i=1}^{k} \frac{\mathcal{E}(X1_{B_i})}{\mathbb{P}(B_i)} 1_{B_i} + b\sum_{i=1}^{k} \frac{\mathcal{E}(Y1_{B_i})}{\mathbb{P}(B_i)} 1_{B_i} \\
&= a\sum_{i=1}^{k} \mathcal{E}(X|\ B_i)1_{B_i} + b\sum_{i=1}^{k} \mathcal{E}(Y|\ B_i)1_{B_i} \\
&= a\mathcal{E}(X|\ \mathcal{P}) + b\mathcal{E}(Y|\ \mathcal{P}).
\end{aligned}$$

2) 我们有

$$\begin{aligned}\mathcal{E}(\mathcal{E}(X|\ \mathcal{P}))&=\mathcal{E}\Big(\sum_{i=1}^{k}\mathcal{E}(X|\ B_i)1_{B_i}\Big)\\&=\sum_{i=1}^{k}\mathcal{E}(X|\ B_i)\mathcal{E}(1_{B_i})\\&=\sum_{i=1}^{k}\mathcal{E}(X|\ B_i)\mathbb{P}(B_i)\\&=\sum_{i=1}^{k}\mathcal{E}(X1_{B_i})\\&=\mathcal{E}\Big(\sum_{i=1}^{k}X1_{B_i}\Big)\\&=\mathcal{E}(X).\end{aligned}$$

3) 令 $Y=\mathcal{E}(X|\ \mathcal{P})$, 根据定义知 Y 是 $\mathcal{P}$ 可测的, 且 (因为$1_{B_i}1_{B_j}=0$,除非$i=j$)

$$\begin{aligned}\mathcal{E}(Y1_{B_i})&=\mathcal{E}(\mathcal{E}(X|\ \mathcal{P})1_{B_i})\\&=\mathcal{E}\Big(\sum_{j=1}^{k}\mathcal{E}(X|\ B_j)1_{B_j}1_{B_i}\Big)\\&=\mathcal{E}(\mathcal{E}(X|\ B_i)1_{B_i})\\&=\mathcal{E}(X|\ B_i)\mathcal{E}(1_{B_i})\\&=\mathcal{E}(X|\ B_i)\mathbb{P}(B_i)\\&=\mathcal{E}(X1_{B_i}).\end{aligned}$$

现在来证明 $Y=\mathcal{E}(X|\ \mathcal{P})$ 是唯一满足条件的随机变量. 假设存在 $\mathcal{P}$ 可测的随机变量 Z, 满足

$$\mathcal{E}(Z1_{B_i})=\mathcal{E}(X1_{B_i}),$$

对 $\mathcal{P}$ 的所有块 B_i. 因为 Z 在 B_i 上是常量, 不妨设 $Z(\omega)=c,\forall\omega\in B_i$. 则 $Z1_{B_i}=X1_{B_i}$, 故

$$\mathcal{E}(Z1_{B_i})=\mathcal{E}(c1_{B_i})=c\mathcal{E}(1_{B_i}).$$

从而

$$c\mathcal{E}(1_{B_i})=\mathcal{E}(X1_{B_i}).$$

因此

$$Z(\omega)=c=\frac{\mathcal{E}(Z1_{B_i})}{\mathcal{E}(1_{B_i})}=\mathcal{E}(X|\ B_i)=\mathcal{E}(X|\ \mathcal{P})(\omega).$$

这表明 $Z = \mathcal{E}(X|\ \mathcal{P})$, 得证.

4) 假设 Z 是 $\mathcal{P}$ 可测的, 不妨设 $Z(\omega) = b, \forall \omega \in B_i$. 那么 $ZX1_{B_i} = bX1_{B_i}$, 因而有

$$\begin{aligned}\mathcal{E}(ZX|\ \mathcal{P})(\omega) &= \mathcal{E}(ZX|\ B_i)\\ &= \frac{\mathcal{E}(ZX1_{B_i})}{\mathbb{P}(B_i)}\\ &= \frac{b\mathcal{E}(X1_{B_i})}{\mathbb{P}(B_i)}\\ &= b\mathcal{E}(X|\ \mathcal{P})(\omega)\\ &= Z(\omega)\mathcal{E}(X|\ \mathcal{P})(\omega).\end{aligned}$$

故 $\mathcal{E}(ZX|\ \mathcal{P}) = Z\mathcal{E}(X|\ \mathcal{P})$, 得证.

5) 在 4) 中取 $X = 1$ 得

$$\mathcal{E}(Z1|\ \mathcal{P}) = Z\mathcal{E}(1|\ \mathcal{P}) = Z.$$

然后用 X 代替 Z 即可.

6) 首先, 对 $\omega \in \Omega$ 有

$$\begin{aligned}\mathcal{E}(\mathcal{E}(X|\ \mathcal{P})|\ \mathcal{Q})(\omega) &= \mathcal{E}(\mathcal{E}(X|\ \mathcal{P})|\ [\omega]_{\mathcal{Q}})\\ &= \frac{\mathcal{E}(\mathcal{E}(X|\ \mathcal{P})1_{[\omega]_{\mathcal{Q}}})}{\mathbb{P}([\omega]_{\mathcal{Q}})}.\end{aligned}$$

现在, 因为 $\mathcal{E}(X|\ \mathcal{P})$ 在 $\mathcal{P}$ 的块上是常数且划分 $\mathcal{Q}$ 是 $\mathcal{P}$ 的细分, 从而 $\mathcal{E}(X|\ \mathcal{P})$ 在 $\mathcal{Q}$ 的每个块上也为常数. 因此

$$\begin{aligned}\mathcal{E}(\mathcal{E}(X|\ \mathcal{P})|\ \mathcal{Q})(\omega) &= \frac{\mathcal{E}(\mathcal{E}(X|\ \mathcal{P})1_{[\omega]_{\mathcal{Q}}})}{\mathbb{P}([\omega]_{\mathcal{Q}})}\\ &= \mathcal{E}(X|\ \mathcal{P})(\omega).\end{aligned}$$

又因为上式对所有的 ω 都成立, 故有

$$\mathcal{E}(\mathcal{E}(X|\ \mathcal{P})|\ \mathcal{Q}) = \mathcal{E}(X|\ \mathcal{P}).$$

下面必须指出

$$\mathcal{E}(\mathcal{E}(X|\ \mathcal{Q})|\ \mathcal{P}) = \mathcal{E}(X|\ \mathcal{P}).$$

可以直接证明, 但是计算有点长. 因此我们利用 3) 来代替直接证明. 首先, 令

$$Y = \mathcal{E}(X|\ \mathcal{Q}).$$

然后根据 3) 知, 对所有的 $B \in \mathcal{A}(\mathcal{Q})$, 有

$$\mathcal{E}(Y1_B) = \mathcal{E}(X1_B). \tag{1}$$

因为 $\mathcal{Q}$ 是 $\mathcal{P}$ 的一个细分, 所以 $\mathcal{A}(\mathcal{P}) \subseteq \mathcal{A}(\mathcal{Q})$, 故上式对所有的 $B \in \mathcal{A}(\mathcal{P})$ 都成立. 现令

$$Z = \mathcal{E}(X|\ \mathcal{P}).$$

然后根据 3) 知, 对所有的 $B \in \mathcal{A}(\mathcal{P})$, 有

$$\mathcal{E}(Z1_B) = \mathcal{E}(X1_B). \tag{2}$$

结合 (1) 和 (2), 得

$$\mathcal{E}(Y1_B) = \mathcal{E}(Z1_B),$$

对所有的 $B \in \mathcal{A}(\mathcal{P})$. 最后, 因为 Z 是 $\mathcal{P}$ 可测的, 由 3) 知

$$Z = \mathcal{E}(Y|\ \mathcal{P}).$$

从而替换 Z 和 Y, 有

$$\mathcal{E}(X|\ \mathcal{P}) = \mathcal{E}(Y = \mathcal{E}(X|\ \mathcal{Q})|\ \mathcal{P}).$$

得证.

7) 假设 $\mathcal{P}_X$ 和 $\mathcal{P}$ 独立, 那么对 $\mathcal{P}$ 的任意块 B, 有

$$\mathbb{P}(X = r|\ B) = \mathbb{P}(X = r).$$

从而

$$\begin{aligned}
\mathcal{E}(X|\ \mathcal{P})(\omega) &= \mathcal{E}(X|\ [\omega]_{\mathcal{P}}) \\
&= \sum_{r \in X(\Omega)} r\mathbb{P}(X = r|\ [\omega]_{\mathcal{P}}) \\
&= \sum_{r \in X(\Omega)} r\mathbb{P}(X = r) \\
&= \mathcal{E}(X).
\end{aligned}$$

得证.

关于随机变量的条件期望

我们可以运用随机变量关于划分的条件期望的结果定义随机变量关于随机变量的条件期望. 确实, 这根本没有什么新的知识 (对有限样本空间来说).

定义 14 设 $(\Omega, \mathbb{P})$ 是一个有限概率空间, Y 是一个离散随机变量其取值为 $\{y_1, \cdots, y_k\}$. 则定义随机变量 X 关于 Y 的条件期望为 X 关于由 Y 生成的划分 $\mathcal{P}_Y$ 的条件期望, 即

$$\mathcal{E}(X|\, Y) = \mathcal{E}(X|\, \mathcal{P}_Y) = \sum_{i=1}^{k} \mathcal{E}(X|\, \{Y = y_i\}) 1_{\{Y=y_i\}}.$$

5.5 随机过程

我们对股票在时间序列点 $t_0 < t_1 < \cdots < t_N$ 上的价格很感兴趣. 如果 t_i 时的股价记为 X_i, 那么初始时这些价格都是未知的, 因此, 可以把它们看作某个概率空间上的一些随机变量. 这就引出了下面一个简单概念, 它在应用数学的许多领域包括金融数学扮演着非常重要的角色.

定义 15 样本空间 Ω 上的 (有限) 随机过程就是 Ω 上的一列随机变量 $X_1, \cdots, X_N$.

随机过程被用来描述现象, 如股票价格过程. 在这些情形中, 生成随机过程的向量存在一定关系. 接下来我们就去挖掘它.

5.6 σ 代数流和鞅

我们利用一个例子来介绍 σ 代数流这个概念.

σ 代数流

考虑如下游戏. 在每个时刻

$$t_0 < t_1 < \cdots < t_N$$

抛掷一枚硬币并记录结果, 形成一个序列.

用 $\{H, T\}^k$ 表示 H 和 T 长度为 k 的所有序列. 这些序列称为 $\{H, T\}$ 上长度为 k 的字符串.

因此, 在时刻 t_i, 游戏的当前状态就是 $\{H, T\}$ 上长度为 i 的一个字符串. 游戏的最终状态构成了 $\{H, T\}$ 上长度为 N 的所有字符串

$$\Omega = \{H, T\}^N = \{e_1 \cdots e_N|\, e_i = H\text{或者}e_i = T\}.$$

状态树

图 7 给出了该游戏状态的一个图示, 称为该游戏的状态树 (图中 $N = 3$). 状态标在树的线段上. t_0 时刻只有一个状态, 没有标出. 它是空字符串.

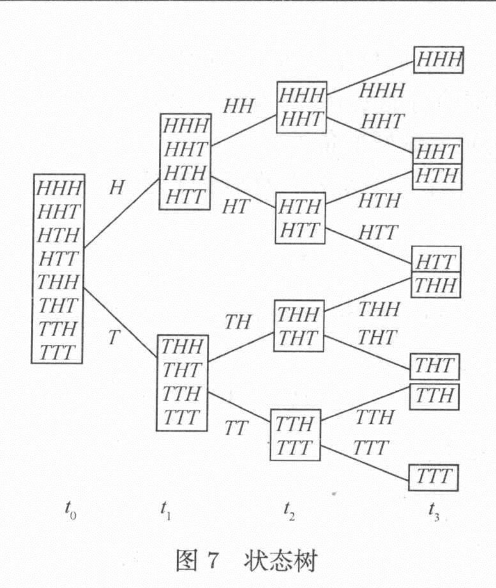


图 7　状态树

方盒子 (树的节点) 包含了在给定当前状态下的所有最后可能状态. 特别地, 如果当前状态是 $\delta = e_1 \cdots e_i$, 那么所有最后可能状态就是前缀为 δ 的所有状态, 用符号表示就是

$$\mathcal{F}_i(\delta) = \{\omega \in \Omega |\ [\omega]_i = \delta\},$$

其中 $[\omega]_i$ 表示 ω 中长度为 i 的前缀. 让我们来观察这些集合.

首先, 在时刻 t_i, 2^i 个子集 $\mathcal{F}_i(\delta)$ 形成了 Ω 的一个划分 $\mathcal{P}_i$. 例如,

$$\mathcal{P}_2 = \{\mathcal{F}_2(HH), \mathcal{F}_2(HT), \mathcal{F}_2(TH), \mathcal{F}_2(TT)\}.$$

其一般情形为

$$\mathcal{P}_i = \{\mathcal{F}_i(\delta_1), \cdots, \mathcal{F}_i(\delta_{2^i})\},$$

其中 $\delta_1, \cdots, \delta_{2^i}$ 是 $\{H, T\}^i$ 中 2^i 个不同元素.

接下来, $\mathcal{P}_i$ 的每个块 $\mathcal{F}_i(\delta)$ 被包含在前面一个划分 $\mathcal{P}_{i-1}$ 的某个块 $\mathcal{F}_{i-1}(\varepsilon)$ 里. 事实上, 有

$$\mathcal{F}_i(\delta) \subseteq \mathcal{F}_{i-1}([\delta]_{i-1}).$$

因此, $\mathcal{P}_i$ 是 $\mathcal{P}_{i-1}$ 的一个细分, 且

$$\mathcal{P}_0 \prec \mathcal{P}_1 \prec \cdots \prec \mathcal{P}_N$$

是 Ω 的一个划分越来越细的序列.

最后, 注意到 $\mathcal{P}_0 = \{\Omega\}$ 是 Ω 最粗糙的可能划分, 只由 Ω 本身这一个块构成. 另一方面, $\mathcal{P}_N = \Omega$ 是最细的可能划分, 因为每个块的长度都为 1, 刚刚包含一个最终状态. 现在我们已经为定义做好准备了.

定义 16　*称 $\Omega = \{\omega_1, \cdots, \omega_m\}$ 上的划分序列, $\mathbb{F} = (\mathcal{P}_0, \cdots, \mathcal{P}_N)$ 是一个 σ 代数流, 如果*

$$\mathcal{P}_0 \prec \mathcal{P}_1 \prec \cdots \prec \mathcal{P}_N,$$

而且, 如果 σ 代数流满足如下条件, 则称它是一个信息结构:

1) $\mathcal{P}_0$ 是最粗糙的可能划分

$$\mathcal{P}_0 = \{\Omega\},$$

表示对 Ω 不了解一点信息.

2) $\mathcal{P}_N$ 是最细的可能划分

$$\mathcal{P}_N = \{\{\omega_1\}, \cdots, \{\omega_m\}\},$$

表示对 Ω 的完全了解, 它的每个块的长度为 1.

因此, 一个信息结构开始对最终状态毫无知情 (除了知道它在 Ω 里), 在每个时刻获得新的信息 (但从来不会丢失信息), 最后完全了解最终状态.

我们应该明确地提到一个 σ 代数流中的划分不需要像例子中那样在长度上双倍增长. 所要求的仅为 $\mathcal{P}_i$ 是 $\mathcal{P}_{i-1}$ 的一个细分.

最后需要注意的是, 在任何时刻 t_i, 在这时可能的状态 δ 和划分 $\mathcal{P}_i$ 之间存在一一对应, 可如下给出

$$\delta \leftrightarrow \{\omega \in \Omega | \ [\omega]_i = \delta\}.$$

这就允许我们将 t_i 时游戏的中间状态和 $\mathcal{P}_i$ 的块 $\mathcal{F}_i(\delta)$ 等同起来. 事实上, 当讨论离散时间衍生品定价模型时, 我们实际上会定义模型的中间状态为一个 σ 代数流的划分的块.

概率

现在假定出现正面的概率为 p, 且每次抛掷硬币是独立的. 那么对任意 $k > 0$, 我们可以在集合 $\{H,T\}^k$ 上定义一个概率测度如下:

$$\mathbb{P}(\delta) = p^{N_H(\delta)} q^{N_T(\delta)},$$

其中

$$N_H(\delta) = \delta\text{中}H\text{的个数},$$
$$N_T(\delta) = \delta\text{中}T\text{的个数}.$$

不难证明 $(\{H,T\}^k, \mathbb{P})$ 是一个有限概率空间.

定理 16 对任意 $\delta \in \{H,T\}^k$, 令

$$\mathbb{P}(\delta) = p^{N_H(\delta)} q^{N_T(\delta)}.$$

1) 如果 $\delta \in \{H,T\}^k$, $\varepsilon \in \{H,T\}^l$, 那么

$$\mathbb{P}(\delta\varepsilon) = \mathbb{P}(\delta)\mathbb{P}(\varepsilon).$$

2) $(\{H,T\}^k, \mathbb{P})$ 是一个概率空间.

证明 关于 1) 有

$$\begin{aligned}\mathbb{P}(\delta\epsilon) &= p^{N_H(\delta\varepsilon)} q^{N_T(\delta\varepsilon)} \\ &= p^{N_H(\delta)+N_H(\varepsilon)} q^{N_T(\delta)+N_T(\varepsilon)} \\ &= p^{N_H(\delta)} p^{N_H(\varepsilon)} q^{N_T(\delta)} q^{N_T(\varepsilon)} \\ &= \mathbb{P}(\delta)\mathbb{P}(\varepsilon).\end{aligned}$$

关于 2), 显然对任意 $\delta \in \{H,T\}^k$ 有

$$0 \leqslant \mathbb{P}(\delta) \leqslant 1.$$

所以只需证明

$$\sum_{\delta \in \{H,T\}^k} \mathbb{P}(\delta) = 1$$

当 $k=1$ 时显然成立, 因为这时有

$$\mathbb{P}(H) + \mathbb{P}(T) = p + (1-p) = 1.$$

对 k 运用归纳法. 假设 k 时是正确的, 那么

$$\begin{aligned}\sum_{\delta \in \{H,T\}^{k+1}} \mathbb{P}(\delta) &= \sum_{\delta \in \{H,T\}^k} \mathbb{P}(H\delta) + \sum_{\delta \in \{H,T\}^k} \mathbb{P}(T\delta) \\ &= \sum_{\delta \in \{H,T\}^k} \mathbb{P}(H)\mathbb{P}(\delta) + \sum_{\delta \in \{H,T\}^k} \mathbb{P}(T)\mathbb{P}(\delta) \\ &= [\mathbb{P}(H) + \mathbb{P}(T)] \sum_{\delta \in \{H,T\}^k} \mathbb{P}(\delta) \\ &= \mathbb{P}(H) + \mathbb{P}(T) \\ &= 1.\end{aligned}$$

所以结果对 $k+1$ 也成立. 因此, 对所有的 $k \geqslant 1$ 都成立.

适应的随机变量

现在假设每出现一次正面, 一个玩家赢 1 美元, 出现一次反面该玩家输 1 美元. 用随机变量 X_i 表示该玩家在 t_i 时的赢利.

因而, t_i 时在状态 $\delta \in \{H,T\}^i$ 下的赢利为

$$N_H(\delta) - N_T(\delta).$$

首先, 看起来很自然地定义 $X_i(\delta)$ 为 $N_H(\delta) - N_T(\delta)$. 问题是这样每个函数 X_i 定义在不同的定义域 $\{H,T\}^i$ 上, 那么函数 (X_i) 就不会形成一个随机过程了.

相反地, 我们在同一个集合 $\Omega = \{H,T\}^N$ 上定义每个 X_i, 而忽略最终状态在 t_i 后的部分. 换句话说, 对任意 $\omega \in \Omega$, 定义

$$X_i(\omega) = N_H([\omega]_i) - N_T([\omega]_i),$$

其中 $[\omega]_i$ 表示 ω 的长度为 i 的前缀. 这样随机变量 X_i 有一个共同的定义域, 但是为了计算 t_i 时的赢利 X_i 不需要游戏状态的“未来信息”.

而且, 这样定义下的函数 X_i 是 $\mathcal{P}_i$ 可测的. 实际上, 对任何 $\omega \in \mathcal{F}_i(\delta)$, 有

$$X_i(\omega) = N_H(\delta) - N_T(\delta).$$

因此, $\mathcal{P}_i$ 的信息暗示了 X_i 的取值信息.

在样本空间 Ω 上有一个 σ 代数流

$$\mathbb{F} = (\mathcal{P}_0, \cdots, \mathcal{P}_N)$$

和一个随机过程

$$\mathbb{X} = (X_0, X_1, \cdots, X_N),$$

其中 $X_0 = 0$, 对所有的 i 有 X_i 是 $\mathcal{P}_i$ 可测的. 由于对所有的 i, 有 X_i 是 $\mathcal{P}_i$ 可测的, 因而我们说随机过程 $\mathbb{X}$ 是 $\mathbb{F}$ 适应的.

鞅

我们想要计算条件期望 $\mathcal{E}(X_{k+1}|\ \mathcal{P}_k)$, 它是在给定 t_k 前面划分的信息下, 赢利在 t_{k+1} 时的期望值. 以此作为结束, 先罗列一些简单的事实.

定理 17　设 $(\{H,T\}^k, \mathbb{P}_k)$ 是一个由如下定义的概率空间:

$$\mathbb{P}_k(\delta) = p^{N_H(\delta)} q^{N_T(\delta)}.$$

那么

1) 对 $\delta \in \{H,T\}^k$, 有

$$\mathbb{P}_N(\mathcal{F}_i(\delta)) = \mathbb{P}_k(\delta).$$

2) 如果 $\delta \in \{H,T\}^k$, $\varepsilon \in \{H,T\}^l$, 那么

$$\mathbb{P}_{k+l+1}(\delta H\varepsilon) \pm \mathbb{P}_{k+l+1}(\delta T\varepsilon) = \mathbb{P}_{k+l}(\delta\varepsilon)(p \pm q).$$

证明 对于 1), 有

$$\mathbb{P}_N(\mathcal{F}_i(\delta)) = \sum_{\sigma \in \{H,T\}^{N-k}} \mathbb{P}_N(\delta\sigma) = \mathbb{P}_k(\delta) \sum_{\sigma \in \{H,T\}^{N-k}} \mathbb{P}_{N-k}(\sigma) = \mathbb{P}_k(\delta).$$

对于 2), 有

$$\begin{aligned}\mathbb{P}_{k+l+1}(\delta H\varepsilon) \pm \mathbb{P}_{k+l+1}(\delta T\varepsilon) &= \mathbb{P}_{k+l+1}(\delta\varepsilon H) \pm \mathbb{P}_{k+l+1}(\delta\varepsilon T)\\ &= \mathbb{P}_{k+l}(\delta\varepsilon)[\mathbb{P}_1(H) \pm \mathbb{P}_1(T)]\\ &= \mathbb{P}_{k+l}(\delta\varepsilon)(p \pm q).\end{aligned}$$

这就完成了证明.

现在可以开始计算 $\mathcal{E}(X_{k+1}|\ \mathcal{P}_k)$ 了. 对任何 $\delta \in \{H,T\}^k$,

$$\begin{aligned}\mathcal{E}(X_{k+1}|\ \mathcal{F}_k(\delta)) &= \sum_{\omega \in \{H,T\}^N} X_{k+1}(\omega)\mathbb{P}_N(\{\omega\}|\ \mathcal{F}_k(\delta))\\ &= \frac{1}{\mathbb{P}_N(\mathcal{F}_k(\delta))} \sum_{\omega \in \{H,T\}^N} X_{k+1}(\omega)\mathbb{P}_N(\{\omega\} \cap \mathcal{F}_k(\delta)).\end{aligned}$$

但是 $\{\omega\} \cap \mathcal{F}_k(\delta)$ 是空集, 除非 $\omega \in \mathcal{F}_k(\delta)$, 这时 $\{\omega\} \cap \mathcal{F}_k(\delta) = \{\omega\}$, 因而

$$\mathcal{E}(X_{k+1}|\ \mathcal{F}_k(\delta)) = \frac{1}{\mathbb{P}_N(\mathcal{F}_k(\delta))} \sum_{\omega \in \mathcal{F}_k(\delta)} X_{k+1}(\omega)\mathbb{P}_N(\{\omega\}).$$

当 ω 取遍集合 $\mathcal{F}_k(\delta)$ 时, 我们可以写成 $\omega = \delta\sigma$, 其中 σ 取遍集合 $\{H,T\}^{N-k}$, 所以

$$\mathcal{E}(X_{k+1}|\ \mathcal{F}_k(\delta)) = \frac{1}{\mathbb{P}_N(\mathcal{F}_k(\delta))} \sum_{\sigma \in \{H,T\}^{N-k}} X_{k+1}(\delta\sigma)\mathbb{P}_N(\delta\sigma).$$

现在, 为了计算 $X_{k+1}(\delta\sigma)$, 需要 $\delta\sigma$ 中长度为 $k+1$ 的前缀, 因此需要知道 σ 中第一个符号的一些信息. 这就促使我们根据 σ 中第一个符号将求和分成两部分, 得

$$\begin{aligned}\mathcal{E}(X_{k+1}|\ \mathcal{F}_k(\delta)) = &\frac{1}{\mathbb{P}_k(\delta)} \sum_{\sigma \in \{H,T\}^{N-k-1}} X_{k+1}(\delta H\sigma)\mathbb{P}_N(\delta H\sigma)\\ &+ \frac{1}{\mathbb{P}_k(\delta)} \sum_{\sigma \in \{H,T\}^{N-k-1}} X_{k+1}(\delta T\sigma)\mathbb{P}_N(\delta T\sigma).\end{aligned}$$

现在可以计算 X_{k+1},

$$\begin{aligned}X_{k+1}(\delta H\sigma) &= N_H(\delta H) - N_T(\delta H)\\ &= 1 + N_H(\delta) - N_T(\delta)\\ &= X_k(\delta) + 1\end{aligned}$$

和

$$\begin{aligned}X_{k+1}(\delta T\sigma) &= N_H(\delta T) - N_T(\delta T)\\ &= N_H(\delta) - N_T(\delta) - 1\\ &= X_k(\delta) - 1.\end{aligned}$$

代入后有

$$\begin{aligned}\mathcal{E}(X_{k+1}|\ \mathcal{F}_k(\delta)) &= \frac{1}{\mathbb{P}_k(\delta)} \sum_{\sigma\in\{H,T\}^{N-k-1}} [X_k(\delta)+1]\mathbb{P}_N(\delta H\sigma)\\ &\quad + \frac{1}{\mathbb{P}_k(\delta)} \sum_{\sigma\in\{H,T\}^{N-k-1}} [X_k(\delta)-1]\mathbb{P}_N(\delta T\sigma)\\ &= \frac{1}{\mathbb{P}_k(\delta)} \sum_{\sigma\in\{H,T\}^{N-k-1}} [\mathbb{P}_N(\delta H\sigma) - \mathbb{P}_N(\delta T\sigma)]\\ &\quad + \frac{X_k(\delta)}{\mathbb{P}_k(\delta)} \sum_{\sigma\in\{H,T\}^{N-k-1}} [\mathbb{P}_N(\delta H\sigma) + \mathbb{P}_N(\delta T\sigma)]\\ &= \frac{1}{\mathbb{P}_k(\delta)} \sum_{\sigma\in\{H,T\}^{N-k-1}} \mathbb{P}_{N-1}(\delta\sigma)(p-q)\\ &\quad + \frac{X_k(\delta)}{\mathbb{P}_k(\delta)} \sum_{\sigma\in\{H,T\}^{N-k-1}} \mathbb{P}_{N-1}(\delta\sigma)(p+q)\\ &= [(p-q)+X_k(\delta)] \sum_{\sigma\in\{H,T\}^{N-k-1}} \mathbb{P}(\sigma)\\ &= (p-q)+X_k(\delta).\end{aligned}$$

因此, 对任何 $\omega\in\Omega$, 我们可以写成 $\omega=\delta\sigma$, 并得到

$$\begin{aligned}\mathcal{E}(X_{k+1}|\ \mathcal{P}_k)(\omega) &= \mathcal{E}(X_{k+1}|\ \mathcal{F}_k(\delta))\\ &= (p-q)+X_k(\delta)\\ &= (p-q)+X_k(\omega).\end{aligned}$$

从而

$$\mathcal{E}(X_{k+1}|\ \mathcal{P}_k) = X_k + (p-q)\mathbf{1},$$

其中 **1** 是取值恒为 1 的随机变量.

$p=q$ 时, 表现出特别重要的意义, 因为这时我们得到

$$\mathcal{E}(X_{k+1}|\ \mathcal{P}_k)=X_k.$$

上式说的是, 如果知道 t_k 时的划分 $\mathcal{P}_k$, 即知道 t_k 时游戏的状态, 那么 t_{k+1} 时赢利的期望等于 t_k 时的赢利. 换句话说, 当 $p=q$ 时我们希望游戏是公平的, 即从一个时刻到下一时刻的期望收入是 0.

我们已经为一个重要的定义做好准备了.

定义 17 称随机过程

$$\mathbb{X}=(X_0,\cdots,X_N)$$

是一个关于 σ 代数流 $\mathbb{F}$ 的鞅 (或 $\mathbb{F}$鞅), 其中

$$\mathbb{F}=(\mathcal{P}_0\prec\mathcal{P}_1\prec\cdots\prec\mathcal{P}_N),$$

如果 $\mathbb{X}$ 是关于 $\mathbb{F}$ 适应的 (即X_i是$\mathcal{P}_i$可测的), 且

$$\mathcal{E}(X_{k+1}|\ \mathcal{P}_k)=X_k.$$

也就是说在给定 $\mathcal{P}_k$ 下, X_{k+1} 的期望刚好是 X_k.

因此, 鞅刻画了公平游戏. 下面的结果表明在给定 $\mathcal{P}_i$ 下, 任何未来随机变量的期望都刚好是 X_i.

定理 18 如果 $\mathbb{X}=(X_0,\cdots,X_N)$ 是一个 $\mathbb{F}$ 鞅, 这里 $\mathbb{F}=(\mathcal{P}_0,\mathcal{P}_1,\cdots,\mathcal{P}_N)$, 那么对任何 $i>0$, 有

$$\mathcal{E}(X_{k+i}|\ \mathcal{P}_k)=X_k.$$

证明 我们知道

$$\mathcal{E}(X_{k+2}|\ \mathcal{P}_{k+1})=X_{k+1}.$$

取条件期望得

$$\mathcal{E}(\mathcal{E}(X_{k+2}|\ \mathcal{P}_{k+1})|\ \mathcal{P}_k)=\mathcal{E}(X_{k+1}|\ \mathcal{P}_k)=X_k.$$

但是根据条件期望的平滑性, 知

$$\mathcal{E}(\mathcal{E}(X_{k+2}|\ \mathcal{P}_{k+1})|\ \mathcal{P}_k)=\mathcal{E}(X_{k+2}|\ \mathcal{P}_k).$$

因此

$$\mathcal{E}(X_{k+2}|\ \mathcal{P}_k)=X_k.$$

通过归纳法就可以完成证明. 我们将此留给读者.

练　习　5

1. 证明：如果 X 是 Ω 上的一个随机变量, 那么

a) $X^{-1}(A \cup B) = X^{-1}(A) \cup X^{-1}(B)$.

b) $X^{-1}(A \cap B) = X^{-1}(A) \cap X^{-1}(B)$.

c) $X^{-1}(A \backslash B) = X^{-1}(A) \backslash X^{-1}(B)$.

2. 设 $\mathcal{P}$ 和 $\mathcal{Q}$ 是非空集 Ω 的两个划分. 证明下列断言等价：

a) $\mathcal{Q}$ 是 $\mathcal{P}$ 的一个细分.

b) $\mathcal{Q}$ 的每个块都是 $\mathcal{P}$ 中块的并.

c) $\mathcal{P}$ 的每个块包含在 $\mathcal{Q}$ 的一个块中.

d) $\mathcal{A}(\mathcal{P}) \subseteq \mathcal{A}(\mathcal{Q})$.

3. 证明：如果 X 是一个非负随机变量, 即对所有的 $\omega \in \Omega$, 有 $X(\omega) \geqslant 0$, 那么 $\mathcal{E}(X|\ \mathcal{P}) \geqslant 0$.

4. 设 X 是 $(\Omega, \mathbb{P})$ 上的随机变量. 证明 X 可以写成由 X 生成的划分的块的示性函数的线性组合.

5. 某个手术使病人完全康复的概率为 60%, 部分康复的概率为 30%, 死亡的概率为 10%. 现考虑一位在手术中生存了下来的病人, 他完全康复的机会有多大?

6. 考虑下面的试验. 有一枚不均匀的硬币, 出现正反面的情况为

$$\mathbb{P}(H) = \frac{2}{3}, \quad \mathbb{P}(T) = \frac{1}{3}.$$

也有两个装有彩色球的缸, 其中

1) 第一个缸里装有 3 个蓝色的球和 5 个红色的球.

2) 第二个缸里装有 7 个蓝色的球和 6 个红色的球.

首先抛掷硬币, 如果出现正面, 从第一个缸里随机地拿出一个球; 如果出现反面, 从第二个缸里随机地拿出一个球. 请问拿到蓝球的概率是多少? 提示：利用全概率公式.

7. 设 Ω 是一个样本空间, $E_1, \cdots, E_n$ 形成 Ω 的一个划分, 且对所有 k, 有 $\mathbb{P}(E_k) \neq 0$. 证明：对 Ω 中的任何事件 A, 有

$$\mathbb{P}(A) = \sum_{k=1}^{n} \mathbb{P}(A|\ E_k)\mathbb{P}(E_k).$$

8. 设 $\mathcal{P} = \{B_1, \cdots, B_k\}$ 是 $(\Omega, \mathbb{P})$ 的一个划分, 且对所有的 i, 有 $\mathbb{P}(B_i) > 0$. 证明贝叶斯公式, 即对 Ω 中任何概率大于零的事件 A, 有

$$\mathbb{P}(E_j|\ A) = \frac{\mathbb{P}(A|\ E_j)\mathbb{P}(E_j)}{\sum_{i=1}^{k} \mathbb{P}(A|\ E_i)\mathbb{P}(E_i)}.$$

9. 证明：任意代数对交和差封闭. 对称差呢 (两个集合的对称差是由两个集合里所有不相交的元素所形成的集合)?

10. 证明：Ω 的任意非空事件类是一个代数当且仅当它包含 Ω, 且对差运算封闭 ($A\backslash B$ 为在 A 而不在 B 中的所有元素集合).

11. 证明：如果事件 $E_1,\cdots,E_n$ 满足 $\mathbb{P}(E_1\cap\cdots\cap E_n)>0$, 那么

$$\mathbb{P}(E_1\cap\cdots\cap E_n)=\mathbb{P}(E_1)\mathbb{P}(E_2|\ E_1)\mathbb{P}(E_3|\ E_1\cap E_2)\cdots\mathbb{P}(E_n|\ E_1\cap\cdots\cap E_{n-1}).$$

12. 证明：对 Ω 的任何划分 $\mathcal{P}$, 集合

$$\mathcal{A}(\mathcal{P})=\{C\subseteq\Omega|\ C=\varnothing\text{或}C=\mathcal{P}\text{中块的并}\}$$

是一个代数.

13. 假设在带有噪音的通讯线 (比如电话线) 上发射由二元表 $\{0,1\}$ 中元素形成的长度为 5 的字符串. 由于存在噪音, 现假定收到一位 (0或1) 的正确概率为 0.75. 假设收到的每位相互独立.

a) 收到一个字符串的正确概率是多少?

b) 收到的一个字符串中刚好有 3 个位是正确的概率为多少?

14. 设 X 和 Y 是 $(\Omega,\mathbb{P})$ 上的随机变量. 假设 X 和 Y 有相同取值 $\{a_1,\cdots,a_n\}$, 且

$$\mathbb{P}(X=a_i)=\mathbb{P}(Y=a_i)=p_i.$$

计算 $\mathbb{P}(X=Y)$.

15. 设 $\mathcal{P}$ 是概率空间 $(\Omega,\mathbb{P})$ 的一个划分. 那么 $\mathcal{E}(\mathbf{1}|\ \mathcal{P})$ 是什么? 其中 $\mathbf{1}$ 是取值恒为 1 的随机变量.

16. 设 $(X_i)_{i=1,\cdots,n}$ 是一个关于 σ 代数流 $(\mathcal{P}_i)_{i=1,\cdots,n}$ 的鞅. 证明对所有的 $k=1,\cdots,n$, 有 $\mathcal{E}(X_k)=\mathcal{E}(X_0)$. 提示：利用 $\mathcal{E}(\mathcal{E}(X|\ \mathcal{P}))=\mathcal{E}(X)$.

17. 设 $X_1,\cdots,X_n$ 是 $(\Omega,\mathbb{P})$ 上有相同均值 μ, 相同取值 $\{r_1,\cdots,r_m\}$ 的随机变量. 令 N 是 $(\Omega,\mathbb{P})$ 上取值 $1,\cdots,n$ 的随机变量. 同时假定 N 与 X_i 独立. 那么可以定义随机变量 S 如下：

$$S=X_1+\cdots+\mathbb{X}_N,$$

这里

$$S(\omega)=X_1(\omega)+\cdots+\mathbb{X}_{N(\omega)}(\omega).$$

请证明：

a) $\mathcal{E}(S|\ N=k)=\mu k$.

b) $\mathcal{E}(S|\ N)=\mu N$.

c) $\mathcal{E}(S) = \mu\mathcal{E}(N)$.

18. 证明：如果 $\mathcal{Q}$ 是 $\mathcal{P}$ 的一个细分, 那么 $\mathcal{E}(\mathcal{E}(X|\ \mathcal{Q})|\ \mathcal{P}) = \mathcal{E}(X|\ \mathcal{P})$.

关于下鞅和上鞅的练习

设 $\mathbb{F} = (\mathcal{P}_0, \cdots, \mathcal{P}_N)$ 是 σ 代数流, $\mathbb{X} = (X_0, \cdots, X_N)$ 是随机过程. 称 $\mathbb{X}$ 是 $\mathbb{F}$ 下鞅, 如果 $\mathbb{X}$ 是 $\mathbb{F}$ 适应的且

$$\mathcal{E}(X_{k+1}|\ \mathcal{P}_k) \geqslant X_k.$$

同样地, 称 $\mathbb{X}$ 是 $\mathbb{F}$ 上鞅, 如果 $\mathbb{X}$ 是 $\mathbb{F}$ 适应的, 且

$$\mathcal{E}(X_{k+1}|\ \mathcal{P}_k) \leqslant X_k.$$

称随机过程 $\mathbb{A} = (A_0, \cdots, A_N)$ 关于 $\mathbb{F}$ 是可料的, 如果对所有的 k, 有 A_k 是 $\mathcal{P}_{k-1}$ 可测的.

19. 证明：$\mathbb{X}$ 是一个 $\mathbb{F}$ 上鞅当且仅当 $-\mathbb{X}$ 是一个 $\mathbb{F}$ 下鞅. 下鞅是公平的赌博吗？它对哪一方有利？上鞅呢？

20. (Doob分解) 设 $(X_0, \cdots, X_N)$ 是一个 $\mathbb{F}$ 适应的随机过程.

i) 证明存在唯一的鞅 $(M_0, \cdots, M_N)$ 和唯一的可料过程 $(A_0, \cdots, A_N)$, 使得 $X_k = M_k + A_k$ 及 $A_0 = 0$. 提示：令 $M_0 = X_0$, $A_0 = 0$.

然后记

$$X_{k+1} - X_k = M_{k+1} - M_k + A_{k+1} - A_k,$$

再关于 $\mathcal{P}_k$ 取条件期望. 利用鞅条件, 将 A_{k+1} 用 A_k 表示出来.

ii) 证明：如果 (X_k) 是一个上鞅, 那么 A_k 是非增的 (即$A_{k+1} \leqslant A_k$). 如果 (X_k) 是一个下鞅呢？

21. 设 $\mathbb{X} = (X_k|\ k = 0, \cdots, T)$ 是一个随机过程.

i) 证明对任何划分 $\mathcal{P}$, 有

$$\max_k\{\mathcal{E}(X_k|\ \mathcal{P})\} \leqslant \mathcal{E}(\max_k\{X_k\}|\ \mathcal{P}).$$

ii) 证明：如果 $\mathbb{X}$ 和 $\mathbb{Y} = (Y_0, \cdots, Y_N)$ 是下鞅, 那么如下定义的 (Z_k) 也是下鞅

$$Z_k = \max\{X_k, Y_k\}.$$

上鞅情形又如何呢？

22. 定义一个随机变量 X 的正部为

$$X^+ = \max\{X, 0\}.$$

如果 $\mathbb{X} = (X_0, \cdots, X_N)$ 是一个鞅, 证明 $\mathbb{X}^+ = (X_0^+, \cdots, X_N^+)$ 是一个下鞅.

第 6 章　离散时间定价模型

我们已经准备好来讨论离散时间定价模型, 即在该模型中, 所有的交易发生在一系列离散时间点上.

衍生品定价问题就是要确定任意衍生品的初始公平价格. 存在的困难是, 在时刻 $t=0$ 不知道衍生品的最终价值, 因为它一般依赖标的物的最终价值. 但是, 可以假定标的物的最终价值为已知的随机变量, 因此我们知道标的物最终价值的可能取值集合. 从而, 也知道了衍生品的最终价值的可能取值集合. 该集合和无套利原理是衍生品定价的关键.

6.1　模型的假设条件

对模型作出如下基本假设.

• 计价物 (numeraire)

所有的价格由一不确定的计价物形式给出. 计价物可能为美元、欧元、英镑、日元等. 诸如 "股票价值 S" 意思是 S 单位计价物. 以后我们会看到, 将模型里的资产的一种作为计价物会很有用. 这样就表示出各资产相对于计价物的价格.

• 无风险资产

假设总有一可行的无风险资产存在. 有关无风险资产的这个想法很简单：在每个时间区间 $[t_{i-1},t_i]$ 内, 无风险资产的价值不会下降而一般会上升. 而且, 在整个区间内的上升值是事先知道的. 在实际当中, 无风险资产的例子有美国国库券和联邦担保存款.

当我们开始探究离散模型时, 由于一些比较直观的原因, 区分资产的价格和资产的数量这两个概念显得较重要, 假设资产的价格随着时间变化, 而资产的数量仅当我们买卖资产时才发生变化. 因此, 一个描述无风险资产的简单方式就是想象有如下表现行为的一种特别资产. 在 t_0 时资产的价格为 1. 在每个区间 $[t_{i-1},t_i]$ 内, 该资产的价格上涨因子为 $\mathrm{e}^{r_i(t_i-t_{i-1})}$, 其中 r_i 为相应区间内的无风险利率.

一般的书都习惯将银行账户或者无风险债券当作模型中的无风险资产. 但是对一个普通的银行账户来说, 不是每单位 (比如说美元) 的价值在变而是数量在变. 假设 t_0 时我们在一个账户中存入 \$10(10 个单位美元), 经过一段 5% 的增长时期后有 \$10.5, 而不是 10 "单位美元" , 每 "单位美元" 值 \$1.05.

不管无风险资产的本质是什么, 我们所分析的重点是资产的价格结构 (在正式定义资产价格时会给出定义).

进一步的假设

除了前面的假设外, 我们还必须做一些不怎么现实的简单假设. 这些假设对我们的分析非常有帮助, 不管它们是否存在, 我们仍能从很大程度上认识到, 基于这些简单模型, 市场是如何起作用的.

• 市场的无限可分

说市场是无限可分的, 意思是我们能说, 比如 $\sqrt{2}$或 $-\pi$ 份股票或债券.

• 市场是无摩擦的

所有的交易能够立即发生而不会有延迟.

• 完美市场

说市场是完美的, 即

没有交易费或佣金,

无卖空限制,

借贷利率一样.

• 买卖平价

作为完美市场概念的一个扩充, 假定任何资产的买价与卖价一样, 也就是说, 如果一种资产能够以价格 S 买进, 那么也能以价格 S 卖出. 例如, 如果能够以每股价格 S 买进一只股票, 那么也能以每股 S 的价格卖出该股票. 如果能够以每股价格 S 买进一只债券, 则相似的债券以 S 出售.

• 在无套利假设下确定价格

如我们已经讨论过的, 如果市场上存在套利机会, 那么价格将会被调整以使得套利机会消失. 因此, 在无套利假设下进行定价是合理的.

6.2 正随机变量

我们将在样本空间 Ω 上定义随机变量的非负性、严格正性和强正性.

定义 1 令 X 是 Ω 上的随机变量. 那么

1) 称 X 是非负的, 记 $X \geqslant 0$, 如果

$$X(\omega) \geqslant 0, \quad \text{对所有的}\omega \in \Omega$$

(也用“positive”来表示随机变量具有该性质).

2) 称 X 是严格正的, 记 $X > 0$, 如果

$$X(\omega) \geqslant 0, \text{对所有的} \omega \in \Omega, \text{且至少存在一个} \omega \in \Omega, \text{使得} X(\omega) > 0.$$

3) 称 X 是强正的, 记 $X \gg 0$, 如果

$$X(\omega) > 0, \text{对所有的} \omega \in \Omega.$$

6.3 举例说明基本模型

在正式定义离散时间模型之前, 通过例子来引导出定义似乎是一个好主意.

假想我们对某一股票很感兴趣, 该股票对利率非常敏感, 当利率下跌时股价一般上涨, 反之一样 (比如房屋建筑类公司股票).

因此, 我们决定采用美国联邦制度理事会使用的折现率来考虑未来时间段内利率的变化. 我们期望一个经济状况对应着一种折现率 (折现率是联邦政府收取银行借钱的利率, 该利率经常被用作其他利率的起始点).

需要着重指出的是, 当建立一个利率模型时, 我们仅需基于经济报告、研究和其他常用工具来推测未来的变化. 但是, 模型必须在 t_0 时建立, 别无其他的选择.

如图 1, 假设当前 t_0 时的折现率为 2%, 同时假设经济环境只有一个状况, 记 ω_0.

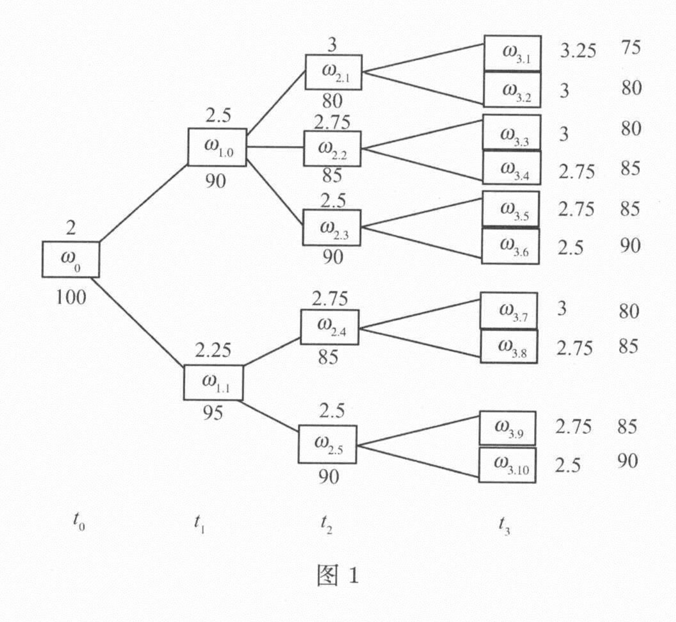


图 1

现在, 根据最新的经济信息, 我们相信在时刻 t_1 联邦政府会提高利率 0.25 个点或者 0.5 个点. 因此, 在时刻 t_1 时经济环境有两个状态, 记 $\omega_{1,0}$, 和$\omega_{1,1}$. 在图 1 中, 利率被标在每个状态上面.

基于好消息我们进一步相信, 在时刻 t_2 联邦政府倾向再次提高利率. 我们推测, 如果前一次利率的增加为 0.5 个点, 这一次的增加量可能为 0.5 或者 0.25 个点, 也可能不变; 如果前一次利率的增加仅为 0.25 个点, 则强烈感觉到这一次的增加量为 0.5 或者 0.25 个点.

一般地, 我们提出一个利率模型, 或者经济的状况, 都是通过观察一段时期内的折现率的路径. 基于预测的利率, 我们同样可推测股票的价格. 在图 1 中, 将股价标在每个状态下面.

注意到, 每一个状态和时间点有一个相应的股价. 因此, 例如, 能如下定义 t_1 时的价格函数 $\hat{S}_1$ 为

$$\widehat{S}_1(\omega_{1,0}) = 90, \quad \widehat{S}_1(\omega_{1,1}) = 95.$$

t_2 时的价格函数 $\widehat{S}_2$ 为

$$\begin{aligned}
\widehat{S}_2(\omega_{2,1}) &= 80, \\
\widehat{S}_2(\omega_{2,2}) &= 85, \\
\widehat{S}_2(\omega_{2,3}) &= 90, \\
\widehat{S}_2(\omega_{2,4}) &= 85, \\
\widehat{S}_2(\omega_{2,5}) &= 90.
\end{aligned}$$

虽然能较容易地理解这些函数, 但当进行数学运算时它们有一个明显的缺点, 就是它们的定义域不一样. 特别地, $\widehat{S}_1$ 是定义在 $\{\omega_{1,0}, \omega_{1,2}\}$ 上, 而 $\widehat{S}_2$ 定义在 $\{\omega_{2,1}, \cdots, \omega_{2,5}\}$ 上.

因而, 比较偏向于使用一列定义在同一概率空间上的价格随机变量. 第一步就是对经济状态采用一种稍微不同的看法. 我们从最终状态集 Ω 开始

$$\Omega = \{\omega_{3,1}, \cdots, \omega_{3,10}\},$$

并定义所有中间状态为最终状态的子集. 图 2 描绘出了该想法.

因此, 比如在时刻 t_1 有两个状态

$$\begin{aligned}
s_1 &= \{\omega_{3,1}, \cdots, \omega_{3,6}\}, \\
s_2 &= \{\omega_{3,7}, \cdots, \omega_{3,10}\}.
\end{aligned}$$

比如, 我们现在就能在 Ω 上定义 t_1 时的价格随机变量 S_1, 赋值 90 给 s_1 中的所有元素, 而赋值 95 给 s_2 中的所有元素. 用符号表示为

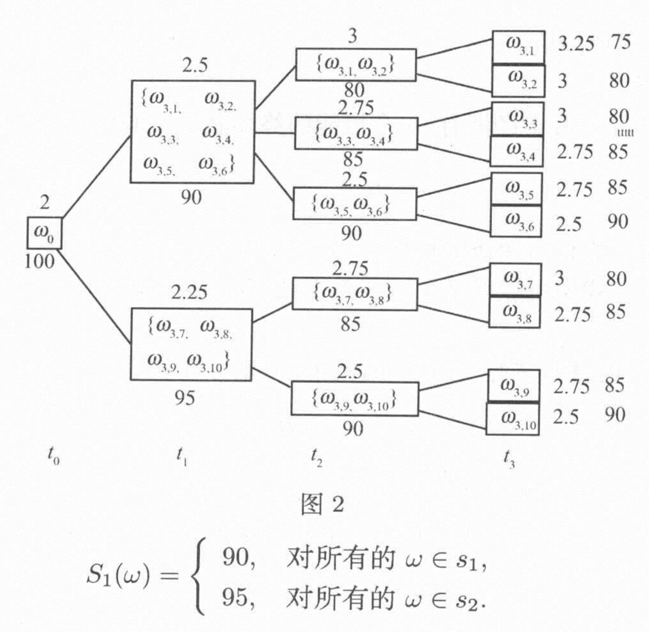


图 2

$$S_1(\omega) = \begin{cases} 90, & \text{对所有的 } \omega \in s_1, \\ 95, & \text{对所有的 } \omega \in s_2. \end{cases}$$

需要强调的是, 该过程只是为了数学上的方便. 谈论一个最终状态在 t_1 时的价格是毫无经济意义的, 因为最终状态在 t_1 时并不存在. 然而, 这种方便处理没有坏处而非常有用.

当然, 为了有意义, 随机变量 S_1 必须在 s_1 和 s_2 的每一元素上都为常数, 我们的例子正是这样.

注意到在每个时间点 t_i, 中间状态集是 Ω 的一个划分 $\mathcal{P}_i$, 且 t_i 时的划分是 t_{i-1} 时的划分的细分. 而且, 价格随机变量 S_i 是 $\mathcal{P}_i$ 可测的.

通过该例子的引导, 我们已经为正式定义一般的离散时间模型做好了准备.

6.4 基本模型

下面为离散时间模型 $\mathbb{M}$ 中的基本要素.

• 时间

模型 $\mathbb{M}$ 有 $T+1$ 个时间点

$$t_0 < t_1 < \cdots < t_T.$$

注意到, 这里刚好有 T 个时间段 $[t_{i-1}, t_i], i = 1, \cdots, T$.

• 资产

模型 $\mathbb{M}$ 有有限种基本资产

$$\mathcal{A} = \{a_1, \cdots, a_n\}.$$

假定资产 a_1 为无风险资产.

● 经济状态

在终端时刻 t_T, 假定经济有 m 个可能的终端状态, 用状态空间 Ω 表示

$$\Omega=\{\omega_1,\cdots,\omega_m\}.$$

初始时 (t_0 时刻), 我们对经济的最终状态一无所知, 除了只知道它落在 Ω 中. 但是, 随着时间的推移, 我们可能得到了一些关于经济的最终可能状态的信息 (而从来不会丢失信息).

为了描述对信息的掌握, 在每个时刻 t_i, 假定存在状态空间 Ω 的一个划分

$$\mathcal{P}_i=\{B_{i,1},\cdots,B_{i,m}\}.$$

称 $\mathcal{P}_i$ 为 t_i 时刻的状态划分. 对任意 $i<T$, $\mathcal{P}_i$ 的块 (block) 对应着 t_i 时经济的可能状态, 称其为中间状态 (intermediate state). 图 3 画出了模型的状态树 (或信息树).

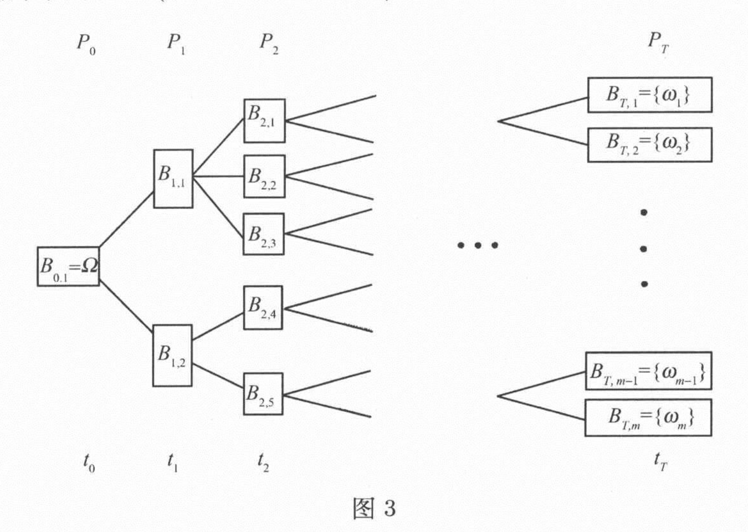


图 3

因此, "状态" 可指中间状态 (它包含了初始状态), 也可指最终状态. 我们也把元素 ω_i 和单点集 $\{\omega_i\}$ 看作一个最终状态, 这样会更加方便.

既然不会丢失信息, 就有 $\mathcal{P}_i$ 是 $\mathcal{P}_{i-1}$ 的一个细分. 事实上, 假定状态流是 Ω 上的一个信息结构

$$\mathbb{F}=(\mathcal{P}_0,\cdots,\mathcal{P}_T).$$

因此

$$\mathcal{P}_0=\{\Omega\}=\{\{\omega_1,\cdots,\omega_n\}\},$$

且

$$\mathcal{P}_T = \{\{\omega_1\}, \cdots, \{\omega_n\}\}.$$

• 客观概率

通常假定在 Ω 上存在一个概率测度, 该测度反映了 Ω 中的每一个最终状态成为最终真实状态的可能性. 称这些可能性为客观概率.

• 资产价格

在离散时间模型中, 每种资产在任何时刻 t_i 不仅有一个价格, 而且该价格可能依赖这时的经济状态. 对每个时刻上的每种资产来说, 就存在一个价格随机变量. 通过前面的例子可以更清楚地认识到, t_i 时的各价格随机变量应该定义在同一样本空间 Ω 上, 且是 $\mathcal{P}_i$可测 的.

定义 2 对每一时刻 t_i 和每一种资产 a_j 来说, 价格随机变量 $S_{i,j}:\Omega \to \mathbb{R}$ 是一个 $\mathcal{P}_i$ 可测的随机变量, $S_{i,j}(\omega)$ 表示最终状态为 ω 时资产 a_j 在时刻 t_i 的价格. 价格随机变量必须满足如下性质:

1) 对无风险资产来说, 价格随机变量为常数, 即不依赖经济状态 (这就是为什么称它为无风险的). 特别地,

$$S_{0,1} = 1,$$

并且对所有的 i, 有

$$S_{i,1} = \mathrm{e}^{r_i(t_i - t_{i-1})} S_{i-1,1},$$

其中 $r_i \geqslant 0$ 为区间 $[t_{i-1}, t_i]$ 上的无风险利率.

2) 对所有其他的资产 $(j > 1)$ 和所有时刻 t_i, 有

$$S_{i,j} \geqslant 0.$$

3) 对固定时刻 t_k, 这时的价格随机向量称为价格向量, 为

$$S_k = (S_{k,1}, \cdots, S_{k,n}).$$

4) 对固定的资产 a_j, 序列

$$(S_{0,j}, \cdots, S_{T,j})$$

是随机过程, 称为资产 a_j 的价格过程. 它描述了整个时期内资产 a_j 的价格变化.

用无风险资产作为计价物

后面我们将会更具体地看到, 使用无风险资产本身作为计价物 (而不是美元、欧元、日元或其他) 将极大地简化问题 (虽然现在看起来并不那么明显).

看一个例子, 假设我们希望评估各种投资的质量. 考虑一个这样的投资, 在 1 年内将 \$100 变成 \$104. 这是一个好的投资吗? 对这个问题我们不好回答, 因为评定

一项投资质量的好坏应相对于某些有保证的标准来说. 例如, 如果无风险资产在 1 年内可将 \$100 变成 \$105, 那么收益 4% 的投资不是好投资.

现在, 如果我们用无风险资产代替美元作为计价物, 那么判断一个投资是否好 (相对于无风险投资) 就非常容易. 例如, 一项投资能够将价值为 100 单位无风险资产变成任何大于 100 单位, 则为一项好的投资, 至少相对于无风险投资的最小标准来说.

资产折现价格 (discounted asset price) 由如下给出

$$\overline{S}_{i,j} = \frac{S_{i,j}}{S_{i,1}}.$$

因此, 折现价格为非折现价格除以无风险资产在同一时间的价格 $S_{i,1}$. 也称 $\overline{S}_{i,j}$ 为 $S_{i,j}$ 的 (0 时期) 现值. 注意到, 特别地有

$$\overline{S}_{i,1} = 1.$$

换句话说, 无风险资产在任何时候的价格都为 1(单位). 这是因为无风险资产相对于自己来说是中性的 (不好也不坏).

折现价格向量为

$$\overline{S}_i = (\overline{S}_{i,1}, \cdots, \overline{S}_{i,n}).$$

6.5　投资组合和交易策略

投资组合用来描述投资者在固定时期内的资产持有情况. 当然, 允许在每一个中间时刻调整资产的持有情况是合理. 允许根据当时的经济状态作出调整也是合理的. 下面为其正式定义.

定义 3　区间 $[t_{i-1}, t_i]$ 上的投资组合为 Ω 上的随机向量

$$\Theta_i = (\theta_{i,1}, \cdots, \theta_{i,n}),$$

这里, $\theta_{i,j}(\omega_k)$ 表示状态为 ω_k 时, 投资者在时刻 t_{i-1} 对资产 a_j 的持有量, 并在 $[t_{i-1}, t_i]$ 内一直持有它. 而且, 要求 $\theta_{i,j}$ 是 $\mathcal{P}_{i-1}$ 可测的. 这对应着一个很显然的事实, 即在 t_{i-1} 时必须知道 $\theta_{i,j}$ 的取值.

值得再提的是: 组合 Θ_i 在时刻 t_{i-1} 被确定, 并被持有至 t_i.

注意到 $\theta_{i,j}$ 在表示数量外也指出了头寸情况: $\theta_{i,j}$ 为正表示多头, $\theta_{i,j}$ 为负表示空头.

较方便的是可定义一个术语 (非标准) 来表示组合中风险资产的持有, 即除了无风险资产外的其他所有资产.

定义 4 在区间 $[t_{i-1},t_i]$ 内的风险持有 (risky holding) 为 Ω 上的随机向量

$$\Theta_i^{\text{risky}} = (\theta_{i,2},\cdots,\theta_{i,n}),$$

其中 $\theta_{i,j}(\omega_k)$ 表示状态为 ω_k 时风险资产 a_j 在时刻 t_{i-1} 的拥有量, 并在 $[t_{i-1},t_i]$ 一直持有. 而且, 要求 $\theta_{i,j}$ 是 $\mathcal{P}_{i-1}$可测 的.

组合的重构

在离散时间模型里, "投资组合" 为如下的动态过程. 在初始时刻 t_0, 投资者的最初投资组合为

$$\Theta_1 = (\theta_{1,1},\cdots,\theta_{1,n}),$$

在整个区间 $[t_0,t_1]$ 内一直持有. 注意到随机变量 $\theta_{1,j}$ 是 $\mathcal{P}_0$ 可测的, 即 $\theta_{1,j}$ 为常数.

到 t_1 时, 投资者必须清算组合 Θ_1, 同时再占有一个新的组合 Θ_2. 当然, 没有谁去阻止该投资者简单地又持有相同的组合, 意思是 $\Theta_2=\Theta_1$. 然而即使是这种情形, 为了一致性可简单地认为是清算后再占有. 因为假设不需要佣金, 这样做并没有害处.

一般而言, 投资者在时刻 t_{i-1} 清算组合 Θ_{i-1}, 并占有新的组合 Θ_i. 称该过程为投资组合的重构. 通过组合重构得到的序列 $\Phi=(\Theta_1,\cdots,\Theta_T)$ 有一个特定的称呼.

定义 5 称如下形式的投资组合序列为模型 $\mathbb{M}$ 的一个交易策略

$$\Phi = (\Theta_1,\cdots,\Theta_T),$$

其中 Θ_i 为区间 $[t_{i-1},t_i]$ 上的投资组合.

我们能从交易策略中分离出单个资产, 而得到用来描述该资产的持有量的随机过程. 特别地, 对每种资产 a_j, 资产持有过程是这样的一个随机过程:

$$\Phi_j = (\theta_{1,j},\cdots,\theta_{T,j}).$$

从而一个交易策略可由随机变量形成的矩阵来表示

$$\Phi = \begin{pmatrix} \Theta_1 \\ \vdots \\ \Theta_T \end{pmatrix} = \begin{pmatrix} \theta_{1,1} & \theta_{1,2} & \cdots & \theta_{1,n} \\ \vdots & \vdots & & \vdots \\ \theta_{T,1} & \theta_{T,2} & \cdots & \theta_{T,n} \end{pmatrix},$$

其中矩阵的行对应着时间 (第一行对应时刻 t_0), 列对应着资产. 事实上, 第 j 列就是资产 a_j 的价格过程.

下面给交易策略的风险部分定义一个 (非标准) 名字.

定义 6　称风险资产持有的序列为模型 $\mathbb{M}$ 的风险子策略 (risky substrategy)

$$\Phi^{\text{risky}} = (\Theta_1^{\text{risky}}, \cdots, \Theta_T^{\text{risky}}),$$

其中 Θ_i^{risky} 为区间 $[t_{i-1}, t_i]$ 上的风险持有.

用矩阵形式表示, 交易策略 Φ 的风险部分由删除矩阵 Φ 第一列后所剩下的其他所有列组成, 即有

$$\Phi^{\text{risky}} = \begin{pmatrix} \Theta_1^{\text{risky}} \\ \vdots \\ \Theta_T^{\text{risky}} \end{pmatrix} = \begin{pmatrix} \theta_{1,2} & \cdots & \theta_{1,n} \\ \vdots & & \vdots \\ \theta_{T,2} & \cdots & \theta_{T,n} \end{pmatrix}.$$

请回忆一下, 称随机过程 $\mathbb{X} = (X_k)$ 关于 $\mathbb{F} = (\mathcal{F}_k)$ 是适应的, 如果对每个 k 都有 X_k 是 $\mathcal{F}_k$ 可测的. 意思是说如果 $\mathcal{F}_k$ 是已知的, 那么 X_k 在时刻 t_k 也是已知的.

另一方面, 一种资产的持有过程 $\Phi_j = (\theta_{1,j}, \cdots, \theta_{T,j})$ 具有性质: $\theta_{k,j}$ 是 $\mathcal{F}_{k-1}$ 可测的. 它对应着这样的事实: 当 $\mathcal{F}_{k-1}$ 是已知时, $\theta_{k,j}$ 在时刻 t_{k-1} 就已知了. 还有许多这样的情况, 例如, 在赌博中下赌注, 玩家在游戏结果出来 (时刻 t_k) 之前就知道了 t_k 时的赌注 X_k. 经常说 t_{k-1} 时就已经知道了 X_k.

定义 7　随机过程

$$\mathbb{X} = (X_1, \cdots, X_T).$$

关于流

$$\mathbb{F} = (\mathcal{P}_0 \prec \mathcal{P}_1 \prec \cdots \prec \mathcal{P}_T)$$

是可料的, 如果对所有的 i, 有 X_i 是 $\mathcal{P}_{i-1}$ 可测的.

运用这种新“语言”, 我们可以说资产持有过程正是可料随机过程的另一称呼. 同样, 交易策略也是可料随机过程.

投资组合的估值

如果 $\Phi = (\Theta_1, \cdots, \Theta_T)$ 是交易策略, 又因为组合 Θ_i 仅在时间段 $[t_{i-1}, t_i]$ 上存在, 故仅在占有时刻 t_{i-1} 和清算时刻 t_i 赋值给 Θ_i 是合理的.

一个投资组合 Θ_i 的占有价值 (acqusition value) 或占有价格由如下内积给出其定义

$$\mathcal{V}_{i-1}(\Theta_i) = \langle \Theta_i, S_{i-1} \rangle = \sum_{j=1}^{n} \theta_{i,j} S_{i-1,j}.$$

组合 Θ_i 的清算价值或清算价格的定义如下

$$\mathcal{V}_i(\Theta_i) = \langle \Theta_i, S_i \rangle = \sum_{j=1}^{n} \theta_{i,j} S_{i,j}.$$

我们也能将投资组合的价值进行折现, 得

$$\overline{\mathcal{V}}_{i-1}(\Theta_i) = \langle \Theta_i, \overline{S}_{i-1} \rangle = \sum_{j=1}^{n} \theta_{i,j} \overline{S}_{i-1,j} = \frac{1}{S_{i-1,1}} \sum_{j=1}^{n} \theta_{i,j} S_{i-1,j}$$

和

$$\overline{\mathcal{V}}_i(\Theta_i) = \langle \Theta_i, \overline{S}_i \rangle = \sum_{j=1}^{n} \theta_{i,j} \overline{S}_{i,j} = \frac{1}{S_{i,1}} \sum_{j=1}^{n} \theta_{i,j} S_{i,j}.$$

注意, 可以直接用折现价格或间接地首先计算非折现价格然后折现而得到折现价值.

自融资交易策略

对一个交易策略 $\Phi = (\Theta_1, \cdots, \Theta_T)$ 来说, 如果 Θ_{i+1} 的占有价格等于 Θ_i 的清算价格, 那么在时刻 t_i 时重建组合既不需要再投入资金也没有资金可拿走.

定义 8 称交易策略 $\Phi = (\Theta_1, \cdots, \Theta_T)$ 是自融资的 (self-financing), 如果在任意时刻 $t_i (i \neq 0, T)$有 Θ_{i+1} 的占有价格等于 Θ_i 的清算价格, 即

$$\mathcal{V}_i(\Theta_{i+1}) = \mathcal{V}_i(\Theta_i).$$

用 $\mathcal{T}$ 表示所有自融资交易策略全体.

因此, 对一个自融资交易策略来说, 最初需投入的资金为 $\mathcal{V}_0(\Theta_1)$, 到 t_T 时进行清算后产生 $\mathcal{V}_T(\Theta_T)$ 这么多的收入. 在交易策略的整个生命周期里, 没有其他资金注入也没有其他资金被抽走.

自融资交易策略集 $\mathcal{T}$ 在如下运算法则下是一个向量空间:

加法

$$(\Theta_{1,1}, \cdots, \Theta_{1,T}) + (\Theta_{2,1}, \cdots, \Theta_{2,T}) = (\Theta_{1,1} + \Theta_{2,1}, \cdots, \Theta_{1,T} + \Theta_{2,T}),$$

点乘

$$a(\Theta_1, \cdots, \Theta_T) = (a\Theta_1, \cdots, a\Theta_T).$$

证明留给读者.

我们可以将符号 $\mathcal{V}_i$ 的用途扩展到自融资交易策略 Φ 上, 作为 Θ_{i+1} 的占有价格和 Θ_i 的清算价格的共同值, 用符号表示为

$$\mathcal{V}_i(\Phi) = \mathcal{V}_i(\Theta_{i+1}) = \mathcal{V}_i(\Theta_i).$$

需要强调的是, 这种扩展只适用于自融资交易策略.

将把 $\mathcal{V}_0(\Phi)$ 看作交易策略 Φ 的初始成本, $\mathcal{V}_T(\Phi)$ 看作交易策略 Φ 的收入. 下一定理给出了估值函数 $\mathcal{V}_i$ 的一些重要性质.

定理 1　1) 对定义在投资组合上的估值函数,

a) 第 i 个估值函数

$$\mathcal{V}_i : RV^n(\Omega) \to RV(\Omega)$$

是一个线性变换, 就是说, 对投资组合 $\Theta_1, \Theta_2 \in RV^n(\Omega)$ 和任意实数 $a, b \in \mathbb{R}$, 有

$$\mathcal{V}_i(a\Theta_1 + b\Theta_2) = a\mathcal{V}_i(\Theta_1) + b\mathcal{V}_i(\Theta_2).$$

b) "占有"随机变量 $\mathcal{V}_i(\Theta_{i+1})$ 是 $\mathcal{P}_i$ 可测的. 换句话说, 在 t_i 时我们就知道了 Θ_{i+1} 的购买价格.

2) 定义在自融资交易策略上的第 i 个估值函数

$$\mathcal{V}_i : \mathcal{T} \to RV(\Omega)$$

是 $\mathcal{T}$ 上的线性变换, 即对任意的 $\Phi_1, \Phi_2 \in \mathcal{T}$ 和任意实数 $a, b \in \mathbb{R}$, 有

$$\mathcal{V}_i(a\Phi_1 + b\Phi_2) = a\mathcal{V}_i(\Phi_1) + b\mathcal{V}_i(\Phi_2).$$

证明留给读者作为练习.

折现收入

可以对自融资交易策略作出关于价格或价值变化的如下定义.

定义 9　设 Φ 是自融资交易策略. 从时刻 t_{i-1} 到 t_i 的折现价格变化为

$$\Delta\overline{S}_i = \overline{S}_i - \overline{S}_{i-1} = (\Delta\overline{S}_{i,1}, \cdots, \Delta\overline{S}_{i,n}).$$

从时刻 t_{i-1} 到 t_i 的折现价值变化为

$$\begin{aligned}\Delta\overline{\mathcal{V}}_i(\Phi) &= \overline{\mathcal{V}}_i(\Phi) - \overline{\mathcal{V}}_{i-1}(\Phi)\\ &= \langle \Theta_i, \Delta\overline{S}_i \rangle\\ &= \langle \Theta_i, (\overline{S}_i - \overline{S}_{i-1}) \rangle\\ &= \sum_{j=1}^{n} \theta_{i,j}(\overline{S}_{i,j} - \overline{S}_{i-1,j}).\end{aligned}$$

折现(累积)收入 $\overline{G}_k$ 为

$$\begin{aligned}\overline{G}_k(\Phi) &= \overline{\mathcal{V}}_k(\Phi) - \overline{\mathcal{V}}_0(\Phi)\\ &= \sum_{i=1}^{k}[\overline{\mathcal{V}}_i(\Phi) - \overline{\mathcal{V}}_{i-1}(\Phi)]\\ &= \sum_{i=1}^{k}\langle\Theta_i, \Delta\overline{S}_i\rangle.\end{aligned}$$

对 $k < l$, 折现收入 $\overline{G}_{k,l}$ 为

$$\begin{aligned}\overline{G}_{k,l}(\Phi) &= \overline{\mathcal{V}}_l(\Phi) - \overline{\mathcal{V}}_k(\Phi)\\ &= \sum_{i=k+1}^{l}[\overline{\mathcal{V}}_i(\Phi) - \overline{\mathcal{V}}_{i-1}(\Phi)]\\ &= \sum_{i=k+1}^{l}\langle\Theta_i, \Delta\overline{S}_i\rangle.\end{aligned}$$

无风险资产的一个重要性质如下. 假设我们拥有一个自融资交易策略 $\Phi = (\Theta_1, \cdots, \Theta_T)$ 的初始价值 $\mathcal{V}_0(\Phi)$, 也知道 Φ 中除了无风险资产 (数量为 $\theta_{i,1}$) 之外的所有其他资产的数量为 $\theta_{i,j}(i = 1, \cdots, T)$. 则自融资条件暗示着无风险资产的数量有且只有一种可能.

从直观上说这是非常合理的. 为了说明这点, 假设 Φ 的初始价值为 \$1000. 如果 Θ_1 中的风险资产的价值为 \$900, 那么无风险资产的量有且只有一种选择, 即为初始价值的剩余部分 $\theta_{1,1} = 100$. 到 t_1 时, 组合 Θ_1 被清算. 假设组合到时的收益为 \$1100. 如果我们知道了风险资产的数量, 因此也知道了组合 Θ_2 中的风险资产价值, 不妨设为 \$1050, 那么无风险资产的数量必须使得它的价值为 \$50. 故数量 $\theta_{2,1} = 50/S_{1,1}$.

一般地, 如果在时刻 t_k 知道 Θ_k 的清算价值 $\mathcal{V}_k(\Theta_k)$ 和风险资产的数量, 因此就知道了风险资产的价值 $\mathcal{V}_k^*(\Theta_{k+1})$, 那么剩余的价值

$$\mathcal{V}_k(\Theta_k) - \mathcal{V}_k^*(\Theta_{k+1})$$

一定全部投资在无风险资产上以使得自融资条件得以满足. 因此

$$\theta_{k+1,1} = \frac{\mathcal{V}_k(\Theta_k) - \mathcal{V}_k^*(\Theta_{k+1})}{S_{k,1}} = \overline{\mathcal{V}}_k(\Theta_k) - \overline{\mathcal{V}}_k^*(\Theta_{k+1}).$$

从上面的讨论可以得到, 有且只有一种自融资交易策略 Φ, 满足对任何给定的

1) 初始价值 $\mathcal{V}_0$, 它是关于 $\mathcal{P}_0$ 可测的随机变量, 也就是说, 是常数随机变量.

2) 风险子策略

$$\Phi^{\text{risky}} = (\Theta_1^{\text{risky}}, \cdots, \Theta_T^{\text{risky}}),$$

即风险资产 $a_2, \cdots, a_m$ 的资产持有过程 $\Theta_2, \cdots, \Theta_m$ 的集合.

用矩阵形式表示, 我们已经指出了初始价值和自融资条件能唯一确定下面矩阵中的缺失值

$$\Phi = \begin{pmatrix} \Theta_1 \\ \vdots \\ \Theta_T \end{pmatrix} = \begin{pmatrix} ? & \theta_{1,2} & \cdots & \theta_{1,n} \\ \vdots & \vdots & & \vdots \\ ? & \theta_{T,2} & \cdots & \theta_{T,n} \end{pmatrix}.$$

因此, 确定一个自融资交易策略所需要的一切就是初始价值和 $n-1$ 个可料过程 (相应于每种风险资产).

锁定收入

假设知道一个自融资交易策略 Φ, 进一步假设我们想在某中间时刻 t_k 锁定折现收入 $\overline{G}_k(\Phi)$. 为了达到目的, 可在时刻 t_k 清算组合 Θ_k, 然后用所得全部收入购买无风险资产. 用数学符号表示为

$$\Theta_{k+1} = (\overline{\mathcal{V}}_k(\Theta_k), 0, \cdots, 0).$$

从这一步开始往前, 交易策略中的各资产数量再也没有改变. 因此新的交易策略 $\Phi' = (\Theta_1', \cdots, \Theta_T')$ 可定义如下

$$\Theta_i' = \begin{cases} \Theta_i, & \text{如果 } i \leqslant k, \\ (\overline{\mathcal{V}}_k(\Theta_k), 0, \cdots, 0), & \text{如果 } i > k. \end{cases}$$

由于在 t_k 到 t_T 这段时间里, 投资组合仅包含无风险资产, 可知折现收入 $\overline{G}_{k,T}(\Phi)$ 为 0, 故

$$\overline{G}_T(\Phi') = \overline{G}_k(\Phi) + \overline{G}_{k,T}(\Phi) = \overline{G}_k(\Phi).$$

这就表明我们锁定了折现收入 $\overline{G}_k(\Phi)$. 将交易策略 Φ' 看作在 t_k 时锁定 Φ 的折现收入而得到的交易策略.

下面对我们所讨论的内容做个总结.

定理 2　1) 折现收入具有可加性, 即对 $j < k < l$, 有

$$\overline{G}_{j,l}(\Phi) = \overline{G}_{j,k}(\Phi) + \overline{G}_{k,l}(\Phi).$$

2) 无风险资产对自融资交易策略的折现收入贡献为 0.

3) 给定任何常数随机变量 V_0 和风险子策略

$$\Phi^{\text{risky}} = (\Theta_1^{\text{risky}}, \cdots, \Theta_T^{\text{risky}}),$$

即由风险资产 $a_2, \cdots, a_m$ 的资产持有过程 $\Phi_2, \cdots, \Phi_m$ 形成的集合. 有且只有一种自融资交易策略满足其初始价值 $\mathcal{V}_0(\Phi) = V_0$ 和风险资产持有过程为给定的 Φ^{risky}. 因此, 确定一个自融资交易策略所需要的一切就是初始价值和 $n-1$ 个可料过程 (对应着每种风险资产).

4) 给定自融资交易策略 Φ 和时刻 t_k, 可以寻找一种自融资交易策略 Φ' 满足在 t_k 时锁定折现收入, 也就是说

$$\overline{G}_T(\Phi') = \overline{G}_k(\Phi) + \overline{G}_{k,T}(\Phi) = \overline{G}_k(\Phi).$$

价值转移

令

$$\Phi = (\Theta_1, \cdots, \Theta_T)$$

是自融资交易策略. 我们考虑这样的问题, 如果初始时改变无风险资产的持有量, 改变量为 $a \in \mathbb{R}$, 那么会发生什么情况呢? 为了继续满足自融资条件, 我们必须在接下来的每个交易时刻都需滚动该资产. 用符号表示新投资组合为

$$\Theta_i' = (\theta_{i,1} + a1_\Omega, \theta_{i,2}, \cdots, \theta_{i,n}) = \Theta_i + a(1_\Omega, 0, \cdots, 0),$$

对 $i = 1, \cdots, T$.

我们花一点时间来看看上述所提及的滚动过程. 例如, 假设初始组合包含 100 份无风险资产, 到 t_1 时清算组合, 从 100 份无风险资产中得到的收入为 $100\mathrm{e}^{r_1(t_1-t_0)}$. 立即用这些资金又买入 100 份无风险资产, 因此, 组合 Θ_2 也刚好包含 100 份无风险资产.

另一方面, 如果初始组合包含 -100 份无风险资产 (空头), 投资者卖出 100 份无风险资产, 其价值为 -100(投资者被 100 份无风险资产 "钩" 住). 到 t_1 时组合被清算, 这时必须花费 $100\mathrm{e}^{r_1(t_1-t_0)}$ 来赎回无风险资产. 然后立即将无风险资产卖出得到资金 $100\mathrm{e}^{r_1(t_1-t_0)}$. 因此, 组合 Θ_2 也刚好包含 100 份空头无风险资产.

对 Φ' 来说, 自融资条件为

$$\mathcal{V}_i(\Theta_i') = \mathcal{V}_i(\Theta_{i+1}').$$

很容易验证上述结论, 我们将细节问题留作练习.

比较两个交易策略 Φ', Φ 的价值, 有

$$\mathcal{V}_i(\Phi') = \mathcal{V}_i(\Phi) + aS_{i,1}.$$

这表明初始时通过增加 a 份无风险资产来改变交易策略的初始价值, 会在整个模型过程中起到影响, 使得从价值上看在 t_i 时对所有的状态都会增加 $aS_{i,1}$. 另一方面, 折现价值的改变量为常数

$$\overline{\mathcal{V}}_i(\Phi') = \overline{\mathcal{V}}_i(\Phi) + a1_\Omega.$$

故折现收入不受影响.

定理 3 令 Φ 为自融资交易策略, $a \in \mathbb{R}$. 通过调整 Φ 中无风险资产的初始持有量, 设改变量为 a, 从而得到自融资交易策略 Φ', 即

$$\Theta_i' = \Theta_i + a(1_\Omega, 0, \cdots, 0),$$

对 $i = 1, \cdots, T$.

1) Φ' 在时刻 t_i 的价值为

$$\mathcal{V}_i(\Phi') = \mathcal{V}_i(\Phi) + aS_{i,1}.$$

特别地有, 初始价值为

$$\mathcal{V}_0(\Phi') = \mathcal{V}_0(\Phi) + a1_\Omega.$$

最终价值为

$$\mathcal{V}_T(\Phi') = \mathcal{V}_T(\Phi) + aS_{T,1}.$$

2) 折现价值为

$$\overline{\mathcal{V}}_i(\Phi') = \overline{\mathcal{V}}_i(\Phi) + a1_\Omega.$$

从而知道折现收入不受价值改变的影响

$$\overline{G}_n(\Phi') = \overline{G}_n(\Phi).$$

3) 特别地, 取 $a = -\mathcal{V}_0(\Phi)$(为常数), 这使得与 Φ 有相同折现收入的自融资策略 Φ' 的初始价值为 0.

6.6 定价问题: 未定权益和复制

我们的目标就是对基本资产的衍生品进行定价. 关于定价, 意思就是在市场是无套利的这个假设下确定衍生品的初始价格.

为了在 0 时刻能有效地对衍生品进行定价，我们需要知道一些关于衍生品在 T 时刻的可能损益的信息. 对股票期权来说，我们已经知道这不是问题. 例如，在两个状态的经济中，假设股票 a_2 的初始成本为 100，最终损益向量为 $(120,90)$，那么执行价为 95 的买权的最终损益向量为 $(25,0)$.

考虑模型中的一种 (或更多) 资产的任意衍生品 $\mathcal{D}$. 该衍生品不是模型的初始部分，但我们想把它加入到模型中去，且使得加入后不会出现套利机会. 我们唯一知道的就是衍生品的最终损益 X，它为 Ω 上的随机变量.

现在从衍生品定价这一点出发，所有关心的事情就是衍生品的损益随机变量 —— 衍生品的具体特征 (买权、卖权、执行价等) 不再重要. 因此，我们真正感兴趣的就是运用无套利来定价 "随机变量". 在本文中，"随机变量" 有个特殊的称呼.

定义 10 称随机变量 $X:\Omega\to\mathbb{R}$ 为或有选择 (alternative) 或未定权益 (contigent claim 也称或有要求权).

注意，有些作者在该定义中要求 X 是非负的，该想法是基于一个期权的要求权不会有负的损益. 在该情形下，"要求权" 将简单到期，但是我们不作这个附加限制.

因此，定价问题就变成了定价未定权益的问题. 对未定权益 X 进行定价的最简单和最直观的方法可能就是寻找一种自融资交易策略 Φ，使其损益向量等于 X，即

$$\mathcal{V}_T(\Phi)=X.$$

然后令 X 的初始价格等于 Φ 的初始价格. 确实，其他任何选择都会导致套利机会的存在. 如果 X 的初始价格 P_0 不等于 $\mathcal{V}_0(\Phi)$，那么投资者可以买进 Φ 和 X 两者中较便宜者，卖出较贵的. 这样马上就能得到净利润，而到最终的时候，投资者通过结算多头而得到的资金刚好能支付空头.

这就引出了下面的定义.

定义 11 令 $X:\Omega\to\mathbb{R}$ 是未定权益. X 的复制交易策略 (replicating trading strategy，或者复制策略 replicating strategy 或对冲策略 hedging strategy) 就是一种自融资交易策略 $\Phi=(\Phi_1,\cdots,\Phi_T)$，其损益等于 X，即

$$\mathcal{V}_T(\Phi)=\mathcal{V}_T(\Theta_T)=X.$$

称未定权益 X 为可达到的 (attainable)，如果至少存在它的一种复制策略. 称模型是完全的，如果每个未定权益都是可达到的.

由所有可达到的未定权益形成的集合 $\mathcal{M}$ 是向量空间 $RV(\Omega)$ 的一个子空间，其中 $RV(\Omega)$ 为 Ω 上的所有随机变量形成的空间. 我们将验证过程留给读者.

对未定权益 X 进行定价的策略是这样的一个复制交易策略过程，首先寻找 X

的复制交易策略 Φ, 然后令 X 的初始价格等于 Φ 的初始价格. 我们一看到下面这个寻找复制交易策略的例子就会知道怎么处理所涉及该策略的问题.

例 1　我们来考虑一个计算可达到的未定权益的复制交易策略的例子. 虽然不是太难, 但是涉及解线性方程组, 一般最好利用计算机来解决该问题.

图4 描述了一个两资产模型的状态树(也包含了股价). 为了计算方便, 假定无风险利率为 0.

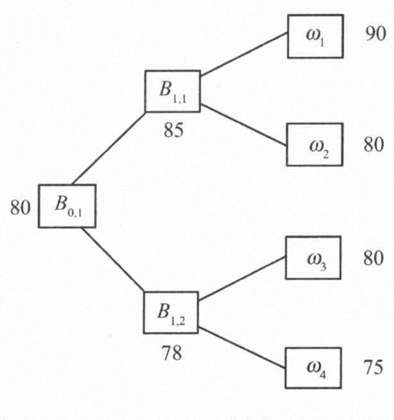


图 4　状态树

我们在后面将看到该模型是完全的, 因此所有未定权益都是可达到的. 我们来计算一种自融资交易策略 $\Phi=(\Theta_1,\Theta_2)$, 它复制了如下未定权益

$$X(\omega_1)=100,\quad X(\omega_2)=90,\quad X(\omega_3)=80,\quad X(\omega_4)=70.$$

也就是, 有 $\mathcal{V}_2(\Theta_2)=X$, 这等价于

$$\begin{aligned}
\mathcal{V}_2(\Theta_2)(\omega_1)&=100,\\
\mathcal{V}_2(\Theta_2)(\omega_2)&=90,\\
\mathcal{V}_2(\Theta_2)(\omega_3)&=80,\\
\mathcal{V}_2(\Theta_2)(\omega_4)&=70.
\end{aligned}$$

展开上述式子得

$$\begin{aligned}
S_{2,1}(\omega_1)\theta_{2,1}(\omega_1)+S_{2,2}(\omega_1)\theta_{2,2}(\omega_1)&=100,\\
S_{2,1}(\omega_2)\theta_{2,1}(\omega_2)+S_{2,2}(\omega_2)\theta_{2,2}(\omega_2)&=90,\\
S_{2,1}(\omega_3)\theta_{2,1}(\omega_3)+S_{2,2}(\omega_3)\theta_{2,2}(\omega_3)&=80,\\
S_{2,1}(\omega_4)\theta_{2,1}(\omega_4)+S_{2,2}(\omega_4)\theta_{2,2}(\omega_4)&=70.
\end{aligned}$$

将实际股价代入得

$$\begin{aligned}
\theta_{2,1}(\omega_1)+90\theta_{2,2}(\omega_1)&=100,\\
\theta_{2,1}(\omega_2)+80\theta_{2,2}(\omega_2)&=90,\\
\theta_{2,1}(\omega_3)+80\theta_{2,2}(\omega_3)&=80,\\
\theta_{2,1}(\omega_4)+75\theta_{2,2}(\omega_4)&=70.
\end{aligned}$$

利用 Θ_2 是 $\mathcal{P}_1$ 可测的, 有

$$\begin{aligned}
\theta_{2,1}(\omega_1)&=\theta_{2,1}(\omega_2),\\
\theta_{2,1}(\omega_3)&=\theta_{2,1}(\omega_4),\\
\theta_{2,2}(\omega_1)&=\theta_{2,2}(\omega_2),\\
\theta_{2,2}(\omega_3)&=\theta_{2,2}(\omega_4).
\end{aligned}$$

因此前面的方程组系统可仅用 ω_1,ω_3 来表示

$$\begin{aligned}
\theta_{2,1}(\omega_1)+90\theta_{2,2}(\omega_1)&=100,\\
\theta_{2,1}(\omega_1)+80\theta_{2,2}(\omega_1)&=90,\\
\theta_{2,1}(\omega_3)+80\theta_{2,2}(\omega_3)&=80,\\
\theta_{2,1}(\omega_3)+75\theta_{2,2}(\omega_3)&=70.
\end{aligned}$$

前面两个方程有唯一解, 后面两个方程也一样, 解之得

$$\begin{aligned}
\Theta_2(\omega_1)=\Theta_2(\omega_2)&=(10,1),\\
\Theta_2(\omega_3)=\Theta_2(\omega_4)&=(-80,2).
\end{aligned}$$

根据时间往后倒推, 我们接下来计算 Θ_2 的占有价值

$$\begin{aligned}
\mathcal{V}_1(\Theta_2)(\omega_1)&=10+85\times 1=95,\\
\mathcal{V}_1(\Theta_2)(\omega_3)&=-80+78\times 2=76.
\end{aligned}$$

自融资条件要求 Θ_1 的清算价值等于 Θ_2 的占有价值, 因此

$$\begin{aligned}
\mathcal{V}_1(\Theta_1)(\omega_1)&=95,\\
\mathcal{V}_1(\Theta_1)(\omega_3)&=76.
\end{aligned}$$

将上述方程展开并代入实际价格, 得

$$\begin{aligned}
\theta_{1,1}(\omega_1)+85\theta_{1,2}(\omega_1)&=95,\\
\theta_{1,1}(\omega_3)+78\theta_{1,2}(\omega_3)&=76.
\end{aligned}$$

但是 Θ_1 是 $\mathcal{P}_0$ 可测的, 即 Θ_1 在 Ω 上是常数, 因此对任意 $\omega \in \Omega$, 有

$$\theta_{1,1}(\omega) + 85\theta_{1,2}(\omega) = 95,$$
$$\theta_{1,1}(\omega) + 78\theta_{1,2}(\omega) = 76.$$

解之得

$$\Theta_1(\omega) = \left(-\frac{950}{7}, \frac{19}{7}\right).$$

在这里, 投资组合由 $950/7 \approx 135.71$ 份债券空头和 $19/7 \approx 2.71$ 份股票多头组成, 其初始价值为

$$-\frac{950}{7} + 80 \times \frac{19}{7} \approx 81.43.$$

因此, 花费 81.43 我们能得到一个投资组合, 该组合保证有如下支付

$$X(\omega_1) = 100, \quad X(\omega_2) = 90, \quad X(\omega_3) = 80, \quad X(\omega_4) = 70.$$

注意, 我们在一些状态获得利润, 在另一些状态则遭受损失. 这就是所期望的无套利模型(随后将证明该模型是无套利的).

一价定律和初始定价泛函

非常明显的是, 复制策略过程只能用来定价可达到的未定权益. 但是, 这仍然有一个潜在的问题, 就是对给定的未定权益存在具有不同初始价值的多个复制策略. 这就需要一价定律来解决该问题了.

定理 4　下面各条件等价:

1) (一价定律) 对所有的交易策略 Φ_1, Φ_2, 有

$$\mathcal{V}_T(\Phi_1) = \mathcal{V}_T(\Phi_2) \Rightarrow \mathcal{V}_0(\Phi_1) = \mathcal{V}_0(\Phi_2).$$

2) 对所有的交易策略 Φ, 有

$$\mathcal{V}_T(\Phi) = 0 \Rightarrow \mathcal{V}_0(\Phi) = 0.$$

证明留给读者.

一价定律保证了下面的初始定价泛函是良定的 (well-defined).

定义 12　在由所有可达到的未定权益形成的向量空间 $\mathcal{M}$ 上, 定义初始定价算子 $\mathcal{I}: \mathcal{M} \to \mathbb{R}$ 如下

$$\mathcal{I}(X) = \mathcal{V}_0(\Phi),\quad \text{对}X\text{的任一复制策略 }\Phi.$$

在离散时间模型中, 对可达到的未定权益定价的关键点就是初始定价泛函的存在性. 因为如果 X 是可达到的未定权益, 也就是说, 如果存在交易策略 Φ, 使得

$$\mathcal{V}_T(\Phi) = X.$$

那么毫不模糊地令 $\mathcal{I}(X) = \mathcal{V}_0(\Phi)$. 而且, 我们能在任何时刻 t_k 定价 X 如下

$$\mathcal{I}_k(X) = \overline{\mathcal{V}}_k(\Phi), \quad \text{对}X\text{的任一复制策略 } \Phi.$$

注意其他任何定价都将导致套利存在. 因为如果在时刻 t_k 有 $\mathcal{I}_k(X) \neq \overline{\mathcal{V}}_k(\Phi)$, 那么投资者在这时就可以进入市场购买 Φ 和 X 两者中较便宜的, 卖出较贵的. 这样在时刻 t_k 就能获得净利润, 而到最终的时候, 投资者通过结算多头而得到的资金刚好能支付空头.

6.7 套利交易策略

现在正式考虑离散时间模型中的套利概念. 想法很简单：套利是这样的情形, 在该情形下没有损失的可能性但有获得收入的可能性. 但是, 在度量相对无风险资产的“自然”保证收入的损失和收益时需要小心. 例如, 假设每年的无风险利率为 10%. 一项 \$100 的投资在 1 年后能得到 \$105, 很难说这是真正的收入, 因为同样多的资金 \$100 投资在无风险资产上能够创造无风险收入 \$110. 因此, 第一项投资相对有保证的无风险投资来说是有损失的.

看起来很自然地去定义套利交易策略 Φ 为这样的交易策略, 其最终的折现收入严格正, 即

$$\overline{G}_T(\Phi) > 0.$$

在一些作者用这个定义的同时, 下面的定义看起来更普遍. 它要求初始价值为 0, 在该情形下折现问题就没有意义了. 需要重点指出的是, 虽然这两个定义不一样, 但在某种意义上是等价的, 看了下面给出的正式定义 (我们所采用的) 后就会更加清楚.

定义 13 *称自融资交易策略 Φ 是套利交易策略 (arbitrage trading strategy 或套利机会 arbitrage opportunity), 如果*

$$\mathcal{V}_0(\Phi) = 0, \quad \text{且 } \mathcal{V}_T(\Phi) > 0.$$

或者等价地用收入表示

$$\mathcal{V}_0(\Phi) = 0, \quad \text{且 } \overline{G}_T(\Phi) > 0.$$

意思是 Φ 的初始成本为0, 但保证在时刻 t_T 不会有损失, 而且至少在一个最终状态下会有正的损益.

我们来证明上述有关套利的两个定义是等价的, 也指出在任何时刻有严格正的折现收入暗示着套利机会的存在. 毕竟, 我们已经知道能锁定一个如此的收入直至到期日.

定理 5 对模型 $\mathbb{M}$ 来说, 下述条件等价:

1) 在 $\mathbb{M}$ 中存在套利机会 Φ, 即

$$\mathcal{V}_0(\Phi) = 0, \quad \text{且 } \overline{G}_T(\Phi) > 0.$$

2) 在 $\mathbb{M}$ 中存在具有严格正的最终折现收入的自融资交易策略 Φ, 即

$$\overline{G}_T(\Phi) > 0.$$

3) 在模型 $\mathbb{M}$ 中存在自融资交易策略 Φ, 该策略在某一时刻 t_k 有严格正的折现收入, 也就是说, 对某个 $k(1 \leqslant k \leqslant T)$,

$$\overline{G}_k(\Phi) > 0.$$

证明 1) $\Rightarrow$ 2) 显然. 通过简单的价值转移可知 2) $\Rightarrow$ 1) (参考定理 3 的最后部分). 2) $\Rightarrow$ 3) 显然. 最后, 如果 3) 成立, 那么我们能够锁定折现收入而得到满足 2) 的交易策略.

6.8 可容许的套利交易策略

一些作者要求套利交易策略从不出现负值, 下面的定义描述了这点.

定义 14 称自融资交易策略Φ是可容许的, 如果在任何时候的价值都非负, 即

$$\mathcal{V}_i(\Phi) \geqslant 0,$$

对所有的 $i = 0, \cdots, T$.

因此, 一个可容许的自融资套利交易策略 Φ 满足:

1) $\mathcal{V}_0(\Phi) = 0$.

2) $\mathcal{V}_i(\Phi) \geqslant 0, \forall i$.

3) $\mathcal{V}_T(\Phi) > 0$.

用收入的形式来表示, 即有

1) $\mathcal{V}_0(\Phi) = 0$.

2) $\overline{G}_i(\Phi) \geqslant 0, \forall i$.

3) $\overline{G}_T(\Phi) > 0$.

下面的结果表明, 要求套利策略是可容许的并不是一个重要的区别.

定理 6 一个模型存在套利机会当且仅当存在可容许的套利机会.

证明 因为一个可容许的套利策略本身就是一个套利策略, 所以只需证明必要性. 也就是说, 在模型中存在一个套利策略

$$\Phi = (\Theta_1, \cdots, \Theta_T).$$

那么也存在一个可容许的套利策略.

当然, 如果 Φ 是可容许的, 那么得证, 因此假设 Φ 不是可容许的. 令 t_k 为存在某状态 $B_{k,u} \in \mathcal{P}_k$, 使得 Φ 的价值为负的最近时间. 既然 $\mathcal{V}_k(\Theta_{k+1})$ 在 $B_{k,u}$ 上为常数, 那么可以记

$$a = \mathcal{V}_k(\Theta_k)(\omega) = \mathcal{V}_k(\Theta_{k+1})(\omega) < 0,$$

对任意 $\omega \in B_{k,u}$.

现在, 问题就变得非常简单了：我们想要隔离产生负的价值的持有, 首先通过令所有其他不相关的价值为 0, 然后利用价值转移来使负的价值这部分变为 0. 下面只需要列出细节.

第一步, 在时刻 t_k 之前我们不做任何处理, 即

$$\Gamma_i = 0, \quad 对 i \leqslant k.$$

从 t_k 往前, 我们采用策略 Φ 当且仅当经济状态在 $B_{k,u}$ 里, Φ 最后一次取负值. 对其他经济状态不做改变. 因此, Γ 的定义为

$$\Gamma_i = \begin{cases} 0, & 如果\ i \leqslant k, \\ 1_{B_{k,u}}\Theta_i, & 如果\ i \geqslant k+1. \end{cases}$$

为了核实这些值, 我们考虑两种情况. 对 $\omega \notin B_{k,u}$, 占有价值和清算价值都总为 0, 即对所有的 i ,

$$\mathcal{V}_i(\Gamma_i)(\omega) = \mathcal{V}_i(\Gamma_{i+1})(\omega) = 0,$$

对 $\omega \in B_{k,u}$, 直到 t_k 时占有价值和清算价值都为 0, 且包括 t_k 时的清算价值. 但是, t_k 时的占有价值为负(等于 a). 随后所有的价值为非负. 因此, 对 $\omega \in B_{k,u}$, 我们可用一种具有启发性的形式来表示价值序列

$$0, \cdots, 0, [\mathcal{V}_k(\Gamma_k)(\omega) = 0, \mathcal{V}_k(\Gamma_{k+1})(\omega) = a < 0], \geqslant 0, \cdots, \geqslant 0.$$

现在, 我们就快接近目标了. 只剩下调整交易策略使其在时刻 t_k 保持自融资条件(在随后的时间也不破坏它). 这可通过仅当状态在 $B_{k,u}$ 中时, 在时刻 t_k 向占有组合 Θ_{k+1} 中加入数量为 $-a/S_{k,1} > 0$ 的无风险资产, 然后滚动该数量来达到.

特别地, 令

$$\Gamma'_i = \begin{cases} 0, & \text{如果} i \leqslant k, \\ 1_{B_{k,u}}\Theta_i - (a1_{B_{k,u}}/S_{k,1}, 0, \cdots, 0), & \text{如果 } i \geqslant k+1. \end{cases}$$

对 $\omega \notin B_{k,u}$, 仍然有

$$\mathcal{V}_i(\Gamma'_i)(\omega) = \mathcal{V}_i(\Gamma'_{i+1})(\omega) = 0.$$

对所有的 i, 但是 $\omega \in B_{k,u}$, 此时的价值为

$$0, \cdots, 0, [\mathcal{V}_k(\Gamma'_k)(\omega) = 0, \mathcal{V}_k(\Gamma_{k+1})(\omega) = 0], \geqslant -a > 0, \cdots, \geqslant -a > 0.$$

因此, $(\mathit{\Gamma}'_1, \cdots, \mathit{\Gamma}'_T)$ 是可容许的自融资交易策略, 定理得证.

6.9　套利的刻画

我们现在来讨论如何刻画套利, 且该刻画也可用来定价未定权益. 这里的关键概念就是鞅测度.

定义 15　设 $\mathbb{M}$ 为离散时间模型. 称 Ω 上的概率分布 Π 为鞅测度 (或者等价鞅测度 equivalent martingale measure, 或风险中性概率测度 risk-neutral probability measure), 如果

1) 概率测度 $\mathbb{P}_\Pi$ 是强正的, 即对所有的 $\omega \in \Omega$, 有

$$\mathbb{P}_\Pi(\omega) > 0.$$

2) 对每一种资产 a_j, 折现价格过程 $(\overline{S}_{0,j}, \cdots, \overline{S}_{T,j})$ 是 $\mathbb{F}$ 鞅, 即对所有的 $k \geqslant 0$, 有

$$\mathcal{E}_\Pi(\overline{S}_{k+1,j}|\mathcal{P}_k) = \overline{S}_{k,j}.$$

或者等价地, 对任意 $i, k \geqslant 0$, 有

$$\mathcal{E}_\Pi(\overline{S}_{k+i,j}|\mathcal{P}_k) = \overline{S}_{k,j}.$$

下一定理根据价值和所得收入给出了鞅测度的刻画.

定理 7　在模型 $\mathbb{M}$ 中, 对强正的概率测度来说, 下面各条件等价:

1) Π 是鞅测度, 就是说, 任何资产的折现价格过程都是一个鞅. 特别地, 对所有的 $k \geqslant 0$,

$$\mathcal{E}_\Pi(\overline{S}_{k+1,j}|\mathcal{P}_k) = \overline{S}_{k,j},$$

或者等价地, 对任意 $i,k \geqslant 0$, 有

$$\mathcal{E}_\Pi(\overline{S}_{k+i,j}|\mathcal{P}_k) = \overline{S}_{k,j}.$$

2) 任何自融资交易策略 Φ 的折现价值过程 $\mathcal{V}_k(\Phi)$ 在 Π 下是一个鞅. 特别地, 对所有的 $k \geqslant 0$,

$$\mathcal{E}_\Pi(\overline{\mathcal{V}}_{k+1}(\Phi)|\mathcal{P}_k) = \overline{\mathcal{V}}_k(\Phi),$$

或者等价地, 对任意 $i,k \geqslant 0$, 有

$$\mathcal{E}_\Pi(\overline{\mathcal{V}}_{k+i}(\Phi)|\mathcal{P}_k) = \overline{\mathcal{V}}_k(\Phi).$$

3) 在任何时候, 任何自融资交易策略 Φ 在 Π 下的期望折现价值等于 Φ 的初始价值, 即对所有的 $k \geqslant 0$, 有

$$\mathcal{E}_\Pi(\overline{\mathcal{V}}_k(\Phi)) = \overline{\mathcal{V}}_0(\Phi),$$

或者等价地, 在 Π 下的期望折现收入为0, 即

$$\mathcal{E}_\Pi(\overline{G}_k(\Phi)) = 0.$$

4) 任何自融资交易策略 Φ 在 Π 下最终收益的折现期望等于 Φ 的初始价值, 即

$$\mathcal{E}_\Pi(\overline{\mathcal{V}}_T(\Phi)) = \overline{\mathcal{V}}_0(\Phi),$$

或者等价地, 在 Π 下最终收入的折现期望等于0, 即

$$\mathcal{E}_\Pi(\overline{G}_T(\Phi)) = 0.$$

进一步, 如果上述任一条件成立, 那么对所有的 $k \geqslant 0$, 有

$$\mathcal{E}_\Pi(\overline{S}_{k,j}) = \overline{S}_{0,j}.$$

意思是说, 资产 a_j 的初始价格等于 a_j 的期望折现价格.

证明 假设 1) 成立, 令 $\Phi = (\Theta_1, \cdots, \Theta_T)$ 是 $\mathbb{M}$ 上的自融资交易策略. 在鞅条件的两边乘以 $\theta_{k+1,j}$, 得

$$\theta_{k+1,j}\mathcal{E}_\Pi(\overline{S}_{k+1,j}|\mathcal{P}_k) = \theta_{k+1,j}\overline{S}_{k,j}.$$

由于 $\theta_{k+1,j}$ 是 $\mathcal{P}_k$ 可测的, 故利用条件期望的性质可将 $\theta_{k+1,j}$ 移到期望算子里面去, 得到

$$\mathcal{E}_\Pi(\theta_{k+1,j}\overline{S}_{k+1,j}|\mathcal{P}_k) = \theta_{k+1,j}\overline{S}_{k,j}.$$

再利用条件期望算子的线性, 对 j 求和有

$$\mathcal{E}_\Pi\left(\sum_{j=1}^{n}\theta_{k+1,j}\overline{S}_{k+1,j}|\mathcal{P}_k\right)=\sum_{j=1}^{n}\theta_{k+1,j}\overline{S}_{k,j}.$$

即有

$$\mathcal{E}_\Pi(\overline{\mathcal{V}}_{k+1}(\Theta_{k+1})|\mathcal{P}_k)=\overline{\mathcal{V}}_k(\Theta_{k+1}),$$

或等价地有

$$\mathcal{E}_\Pi(\overline{\mathcal{V}}_{k+1}(\Phi)|\mathcal{P}_k)=\overline{\mathcal{V}}_k(\Phi).$$

这就是 $\overline{\mathcal{V}}_k(\Phi)$ 为鞅的一个充分条件, 故 2) 成立.

假设 2) 成立, 那么

$$\mathcal{E}_\Pi(\overline{\mathcal{V}}_{k+i}(\Phi)|\mathcal{P}_k)=\overline{\mathcal{V}}_k(\Phi).$$

令 $k=0$, 得

$$\mathcal{E}_\Pi(\overline{\mathcal{V}}_i(\Phi)|\mathcal{P}_0)=\overline{\mathcal{V}}_0(\Phi),$$

或者

$$\mathcal{E}_\Pi(\overline{\mathcal{V}}_i(\Phi))=\overline{\mathcal{V}}_0(\Phi),$$

或者用收入的形式表示, 有

$$\mathcal{E}_\Pi(\overline{G}_i(\Phi))=\mathcal{E}_\Pi(\overline{\mathcal{V}}_i(\Phi)-\overline{\mathcal{V}}_0(\Phi))=0.$$

这就得到了 3) 是成立的. 当然 3) 暗含着 4), 因为后者仅是前者的一种特殊情况.

现在假设 4) 成立, 故对所有的自融资交易策略 Φ, 有

$$\mathcal{E}_\Pi(\overline{G}_T(\Phi))=0.$$

考虑如下交易策略, 固定 $\mathcal{P}_k$ 的划分块 $B_{k,u}$, 也固定资产 a_j. 在时间 t_k 之前不拥有任何东西, 即

$$\Theta_1=\Theta_2=\cdots=\Theta_k=0.$$

在时刻 t_k, 仅当状态在 $B_{k,u}$ 里时才采取行动, 在这种情况下借入一份资产 a_j 所需的成本 $S_{k,j}$, 然后买进该资产一份. 因为债券在 t_k 时价值 $S_{k,1}$, 故 $-S_{k,j}$ 美元的现金等价于 $-S_{k,j}/S_{k,1}=-\overline{S}_{k,j}$ 份债券. 因此, 在时刻 t_k 时的投资组合为

$$\Theta_{k+1}=(-\overline{S}_{k,j}1_{B_{k,u}},0,\cdots,0,1_{B_{k,u}},0,\cdots,0).$$

该组合的占有价值为

$$\mathcal{V}_k(\Theta_{k+1})=(-\overline{S}_{k,j}1_{B_{k,u}})S_{k,1}+1_{B_{k,u}}S_{k,j}=0=\mathcal{V}_k(\Theta_k).$$

因此, 在时刻 t_k, Φ 确实满足自融资条件.

t_{k+1} 时, 在清算 Θ_{k+1} 后仅仅投资无风险资产. 然后一直滚动该资产直到模型的最后. 因此, $\Theta_{k+2},\cdots,\Theta_{k+T}$ 只包含无风险资产, 故从 t_{k+1} 开始往前就没有折现收入.

唯一的折现收入发生在区间 $[t_k,t_{k+1}]$ 内. 因此, 自融资交易策略 $\Phi=(\Theta_1,\cdots,\Theta_T)$ 的折现收入为

$$\begin{aligned}
\overline{G}_T(\Phi)&=\overline{G}_{0,k}(\Phi)+\overline{G}_{k,k+1}(\Phi)+\overline{G}_{k+1,T}(\Phi)\\
&=\overline{G}_{k,k+1}(\Phi)\\
&=\overline{\mathcal{V}}_{k+1}(\Theta_{k+1})-\overline{\mathcal{V}}_k(\Theta_k)\\
&=\overline{\mathcal{V}}_{k+1}(\Theta_{k+1})\\
&=(-\overline{S}_{k,j}1_{B_{k,u}})\overline{S}_{k+1,1}+1_{B_{k,u}}\overline{S}_{k+1,j}\\
&=-\overline{S}_{k,j}1_{B_{k,u}}+1_{B_{k,u}}\overline{S}_{k+1,j}.
\end{aligned}$$

由假设知上式的期望值为0, 故

$$\mathcal{E}_\Pi(-\overline{S}_{k,j}1_{B_{k,u}}+\overline{S}_{k+1,j}1_{B_{k,u}})=0,$$

或

$$\mathcal{E}_\Pi(\overline{S}_{k,j}1_{B_{k,u}})=\mathcal{E}_\Pi(\overline{S}_{k+1,j}1_{B_{k,u}}).$$

两边同时除以 $B_{k,u}$ 的概率, 得

$$\mathcal{E}_\Pi(\overline{S}_{k,j}|B_{k,u})=\mathcal{E}_\Pi(\overline{S}_{k+1,j}|B_{k,u}).$$

因为上式对任何 u 都成立, 故

$$\mathcal{E}_\Pi(\overline{S}_{k,j}|\mathcal{P}_k)=\mathcal{E}_\Pi(\overline{S}_{k+1,j}|\mathcal{P}_k).$$

但是 $\overline{S}_{k,j}$ 是 $\mathcal{P}_k$ 可测的, 从而有

$$\mathcal{E}_\Pi(\overline{S}_{k+1,j}|\mathcal{P}_k) = \overline{S}_{k,j}.$$

这刚好是折现资产定价过程的鞅条件, 因此 Π 是鞅测度且 1) 成立. 定理得证.

资产定价第一基本定理

前一定理清晰地表明鞅测度是我们想要的. 资产定价第一基本定理准确地告诉我们什么时候存在鞅测度.

定理 8(资产定价第一基本定理)　对一个离散时间模型 $\mathbb{M}$ 来说, 如下条件等价:

1) 没有套利交易策略.

2) 在 $\mathbb{M}$ 中存在鞅测度 Π.

证明　证明该定理的要点就是定理 5 中的套利刻画和定理 7 中的鞅测度的刻画, 通过凸分析理论中的一个事实来连接它们.

我们先来比较一下这两个性质. 因为套利暗示着存在自融资交易策略 Φ 满足 $\overline{G}_T(\Phi) > 0$, 因此有

无套利: 对所有的自融资交易策略 Φ, 有

$$\overline{G}_T(\Phi) \not> 0.$$

鞅测度 Π 的存在性: 对所有的自融资交易策略 Φ, 有

$$\mathcal{E}_\Pi(\overline{G}_T(\Phi)) = 0.$$

我们必须指出上述两个性质是等价的. 实际上一方面非常容易. 如果在 $\mathbb{M}$ 中存在鞅测度 Π, 那么不可能存在自融资交易策略 Φ 满足

$$\overline{G}_T(\Phi) > 0.$$

假如存在, 这样的交易策略在一个强正的概率测度下的期望一定为正. 为了看清这一点, 假设

$$\Pi = (\pi_1, \cdots, \pi_m),$$

其中对所有的 i 有 $\pi_i > 0$. 并且假设

$$\overline{G}_T(\Phi) = (g_1, \cdots, g_m),$$

其中对所有的 i, 有 $g_i \geqslant 0$, 且对某些 j 有 $g_j > 0$. 从而

$$\mathcal{E}_\Pi(\overline{G}_T(\Phi)) = \pi_1 g_1 + \cdots + \pi_m g_m > 0.$$

因此这违反了鞅测度条件, 故不存在套利机会. 这就证明了第一基本定理的一方面.

另一方面, 我们必须指出无套利意味着存在鞅测度. 为了让证明更具有几何风格, 我们将一个随机变量看作一个向量而不是一个函数. 样本空间 Ω 的有限保证了这样做的可行性. 特别地, 如果将 Ω 中元素的顺序固定, 比如说 $\Omega = (\omega_1, \cdots, \omega_m)$, 那么任何随机变量 $X : \Omega \to \mathbb{R}$ 就能和它的取值向量等同起来, 取值向量如下

$$\overrightarrow{X} = (X(\omega_1), \cdots, X(\omega_m)).$$

非常明显, 随机变量 X 非负, 严格正或强正当且仅当相应的向量也有该性质.

现在, Π 为鞅测度的条件就能写成下面用内积表示的条件

$$\langle \overrightarrow{\overline{G}_T(\Phi)}, \Pi \rangle = 0.$$

这是因为

$$\langle \overrightarrow{\overline{G}_T(\Phi)}, \Pi \rangle = \sum_{i=1}^{m} \overline{G}_T(\Phi)(\omega_i)\pi_i = \mathcal{E}_\Pi(\overline{G}_T(\Phi)).$$

我们考虑所有最终收入向量形成的集合 $\mathcal{G}$

$$\mathcal{G} = \{\overrightarrow{\overline{G}_T(\Phi)} | \Phi\text{为交易策略}\} \subseteq \mathbb{R}^m.$$

由于价值算子 $\overline{\mathcal{V}}_T$ 和 $\overline{\mathcal{V}}_0$ 都是线性变换, 因此 $\overline{G}_T$ 和 $\overline{G}_T$ 的象 $\mathcal{G}$(为 $\mathbb{R}^m$ 的子空间) 也是线性变换. 无套利条件

$$\overline{G}_T(\Phi) \not> 0$$

等价于除了原点, 向量空间 $\mathcal{G}$ 与 $\mathbb{R}^n$ 中的非负向量集 $\mathbb{R}^n_+$ 不相交,

$$\mathbb{R}^n_+ = \{(x_1, \cdots, x_n) | x_i \geqslant 0\}.$$

即

$$\mathcal{G} \cap \mathbb{R}^n_+ = \{0\}.$$

从附录 B 的定理 5 知 $\mathcal{G}^\perp$ 包含强正的向量 $\Pi = (\pi_1, \cdots, \pi_m)$. 换句话说, 对任何自融资交易策略 Φ, 都有

$$\langle \overrightarrow{\overline{G}_T(\Phi)}, \Pi \rangle = 0.$$

这表明 Π 是鞅测度. 定理得证.

资产定价第二基本定理

现在我们将注意力放到资产定价第二基本定理上来. 回忆一下, 说模型 $\mathbb{M}$ 是完全的, 如果 $\mathbb{R}^m$ 中的每一个未定权益都是可达到的. 也就是说, 如果对每一个 $X \in \mathbb{R}^m$, 存在自融资交易策略 Φ, 使得

$$\mathcal{V}_T(\Phi) = X.$$

接下来我们将利用线性代数里的一个结论. Ω 上的任何强正 (strongly positive) 概率分布 $\varGamma = (\gamma_1, \cdots, \gamma_m)$, 这里 $\mathbb{P}_\varGamma(\omega_k) = \gamma_k$, 在向量空间 $\mathbb{R}^m$ 上定义内积如下

$$\langle X, Y \rangle_\varGamma = \sum_{i=1}^{m} x_i y_i \gamma_i.$$

请读者验证上述的定义确实满足内积性质, 即

1) (双线性)

$$\langle aX + bY, Z \rangle_\varGamma = a\langle X, Z \rangle_\varGamma + b\langle Y, Z \rangle_\varGamma,$$
$$\langle X, aY + bZ \rangle_\varGamma = a\langle X, Y \rangle_\varGamma + b\langle X, Z \rangle_\varGamma.$$

2) (对称性)

$$\langle X, Y \rangle_\varGamma = \langle Y, X \rangle_\varGamma.$$

3) (正定性)

$$\langle X, X \rangle_\varGamma = 0,$$

取等号当且仅当 $X = 0$.

观察到如果记 $\mathbf{1} = (1, \cdots, 1)$, 那么对任意向量 (随机变量)$X$, 有

$$\langle X, \mathbf{1} \rangle_\varGamma = \sum_{i=1}^{m} x_i \gamma_i = \varepsilon_\varGamma(X).$$

我们现在就来看看资产定价第二基本定理.

定理 9 (资产定价第二基本定理) *令 $\mathbb{M}$ 为一个无套利机会的模型. 那么在 $\mathbb{M}$ 上存在唯一的鞅测度当且仅当模型 $\mathbb{M}$ 是完全的.*

证明 我们首先证明 $\mathbb{M}$ 的完全性意味着 $\mathbb{M}$ 上的鞅测度是唯一的. 假设模型 $\mathbb{M}$ 是完全的, Π_1 和 Π_2 是其上的两个鞅测度. 我们想要证明的是 $\Pi_1 = \Pi_2$.

因为 Π_1 是鞅测度, 据定理 7 有

$$\mathcal{E}_{\Pi_1}(\overline{\mathcal{V}}_T(\Phi)) = \overline{\mathcal{V}}_0(\Phi).$$

同样地有

$$\mathcal{E}_{\Pi_2}(\overline{\mathcal{V}}_T(\Phi)) = \overline{\mathcal{V}}_0(\Phi).$$

因此

$$\mathcal{E}_{\Pi_1}(\overline{\mathcal{V}}_T(\Phi)) = \mathcal{E}_{\Pi_2}(\overline{\mathcal{V}}_T(\Phi)).$$

又由于折现期是一样的, 有

$$\mathcal{E}_{\Pi_1}(\mathcal{V}_T(\Phi)) = \mathcal{E}_{\Pi_2}(\mathcal{V}_T(\Phi)).$$

但是由于 $\mathbb{M}$ 是完全的, 故对某一自融资交易策略来说, Ω 上的所有随机变量均有 $\mathcal{V}_T(\Phi)$ 这样的形式. 因此对 Ω 上的所有随机变量 X, 有

$$\mathcal{E}_{\Pi_1}(X) = \mathcal{E}_{\Pi_2}(X).$$

对 $\omega \in \Omega$, 取 $X = 1_{\{\omega\}}$. 从而得到

$$\mathbb{P}_{\Pi_1}(\omega) = \mathbb{P}_{\Pi_2}(\omega).$$

这就意味着 $\Pi_1 = \Pi_2$. 因此 $\mathbb{M}$ 上的鞅测度是唯一的.

为了证明另一方面, 假设 Π 是 $\mathbb{M}$ 上的鞅测度, 且市场是不完全的. 我们想要找到 $\mathbb{M}$ 上的另一不同鞅测度 Π^*. 像证明资产定价第一基本定理那样, 将 Ω 中元素的顺序固定, 比如说 $\Omega = (\omega_1, \cdots, \omega_m)$, Ω 上的随机变量看作 $\mathbb{R}^m$ 中的向量.

因为 $\mathbb{M}$ 是不完全的, 所以存在不可达到的向量(未定权益). 换句话说, 由所有可到达的向量形成的向量空间 $\mathcal{M}$ 是 $\mathbb{R}^m$ 的真子空间.

考虑利用鞅测度 Π 在 $\mathbb{R}^m$ 上定义的内积

$$\langle X, Y \rangle_\Pi = \sum_{i=1}^m x_i y_i \pi_i.$$

线性代数里有个简单的结果是, 如果 $\mathbb{R}^m$ 的一个子空间, 比如 $\mathcal{M}$, 不等于 $\mathbb{R}^m$, 那么存在一个向量 $z = (z_1, \cdots, z_m)$ 与该子空间里的所有向量正交 (垂直). 因此, 对任何可达到的向量 $X = (x_1, \cdots, x_m)$, 有

$$\langle X, Z \rangle_\Pi = \sum_{i=1}^m x_i z_i \pi_i = 0.$$

而且, 因为向量 $\mathbf{1} = (1, \cdots, 1)$ 是可达到的(只需买 $1/S_{T,1}$ 份无风险资产, 然后滚动下去), 故

$$0 = \langle \mathbf{1}, Z \rangle_\Pi = \sum_{i=1}^m z_i \pi_i = \mathcal{E}_\Pi(Z).$$

现在我们试图在 $\mathbb{M}$ 上定义另一不同的鞅测度 $\Pi^* = (\pi_1*, \cdots, \pi_m^*)$. 该概率测度必须是强正的, 必须满足鞅条件, 且不同于 Π.

当然, 首先它必须是一个概率测度. 注意到

$$\sum_{i=1}^{m} z_i \pi_i = 0.$$

我们可尝试如下形式

$$\pi_i^* = \pi_i + c z_i \pi_i,$$

这里 c 为常数. 至少按上面构造的是一个概率测度

$$\sum_{i=1}^{m} \pi_i^* = \sum_{i=1}^{m} \pi_i + c \sum_{i=1}^{m} z_i \pi_i = \sum_{i=1}^{m} \pi_i = 1.$$

另外, 由于 $Z \perp \mathcal{M}$, 从而对任意可达到的向量 $X \in \mathcal{M}$, 有

$$\begin{aligned}
\mathcal{E}_{\Pi^*}(X) &= \sum_{i=1}^{m} x_i \pi_i^* \\
&= \sum_{i=1}^{m} x_i (\pi_i + c z_i \pi_i) \\
&= \sum_{i=1}^{m} x_i \pi_i + c \sum_{i=1}^{m} x_i z_i \pi_i \\
&= \mathcal{E}_{\Pi}(X) + \langle X, Z \rangle_{\Pi} \\
&= \mathcal{E}_{\Pi}(X).
\end{aligned}$$

因此, 对任何自融资交易策略 Φ, 有

$$\mathcal{E}_{\Pi^*}(\overline{\mathcal{V}}_T(\Phi)) = \mathcal{E}_{\Pi}(\overline{\mathcal{V}}_T(\Phi)).$$

又因为 Π 是一个鞅测度, 从而据定理 7 可推得

$$\mathcal{E}_{\Pi^*}(\overline{\mathcal{V}}_T(\Phi)) = \mathcal{E}_{\Pi}(\overline{\mathcal{V}}_T(\Phi)) = \overline{\mathcal{V}_0}(\Phi).$$

但是定理 7 告诉我们 Π^* 也是一个鞅测度, 如果 Π^* 是强正的. 因此, 为了完成该定理的证明, 我们需要做的就只需选一个常数 c, 使得 Π^* 是强正的, 即对所有的 i, 有

$$\pi_i^* = \pi_i + c z_i \pi_i > 0,$$

或等价地

$$1 + c z_i > 0.$$

最后, 令 $M=\max_i\{|z_i|\}$. 那么

$$-M\leqslant z_i\leqslant M.$$

故有

$$-1\leqslant\frac{z_i}{M}\leqslant 1.$$

同时除以 2 再加上 1, 得

$$\frac{1}{2}\leqslant 1+\frac{z_i}{2M}\leqslant\frac{3}{2}.$$

因此可取

$$c=\frac{1}{2M}.$$

这就完成了证明.

6.10 求解鞅测度

现在考虑如何求解一个模型 $\mathbb{M}$ 的鞅测度

$$\Pi=(\pi_1,\cdots,\pi_m).$$

虽然把求解细节过程写下来有点杂乱冗长, 但是技巧很简单.

首先, 注意到任意最终结果 $\omega_r\in\Omega$ 都在一列块 (block) 里面, 列中的每个块来自于每个划分 $\mathcal{P}_k$, 不妨说

$$\{\omega_r\}=B_{T,i_T}\subseteq B_{T-1,i_{T-1}}\subseteq\cdots\subseteq B_{0,i_0}=\Omega.$$

从而 $\pi_r=\mathbb{P}_\Pi(\omega_r)$ 刚好是条件概率的乘积

$$\begin{aligned}\pi_r&=\mathbb{P}_\Pi(\omega_r)\\&=\mathbb{P}_\Pi(\omega_r|B_{T-1,i_{T-1}})\mathbb{P}_\Pi(B_{T-1,i_{T-1}})\\&=\mathbb{P}_\Pi(\omega_r|B_{T-1,i_{T-1}})\mathbb{P}_\Pi(B_{T-1,i_{T-1}}|B_{T-2,i_{T-2}})\mathbb{P}_\Pi(B_{T-2,i_{T-2}})\\&\qquad\cdots\cdots\cdots\cdots\\&=\mathbb{P}_\Pi(\omega_r|B_{T-1,i_{T-1}})\mathbb{P}_\Pi(B_{T-1,i_{T-1}}|B_{T-2,i_{T-2}})\cdots\mathbb{P}_\Pi(B_{1,i_1}|B_{0,i_0}).\end{aligned}$$

因此, 我们就能计算 Π 中的概率值, 如果能计算下面的条件概率

$$\mathbb{P}_\Pi(B_{k+1,u}|B_{k,v}),\tag{6.10.1}$$

对所有的块对 $B_{k+1,u} \subseteq B_{k,v}$.

状态信息树给出了条件概率一个非常直观的图像, 以及它们如何合并在一起来得到鞅测度的. 图 5 指出了从初始块 $B_{0,1}$ 到最终块 $\omega_r = B_{4,i_4}$ 的一条路径. 经常把条件概率标在路径的边上.

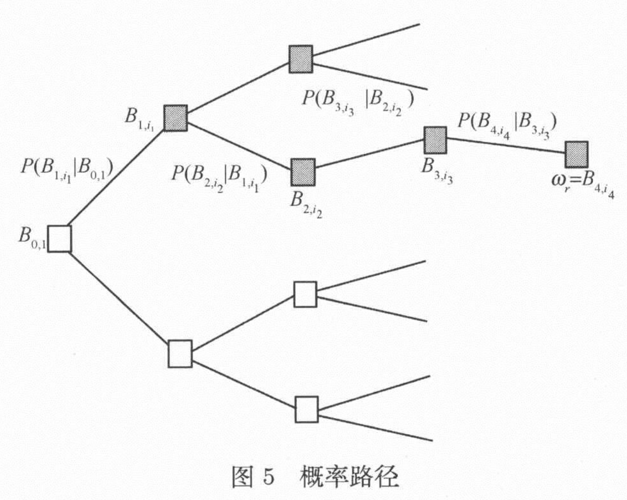


图 5 概率路径

而且, 鞅概率 $\pi_r = \mathbb{P}_\Pi(\omega_r)$ 刚好为标在从 $B_{0,1}$ 到 ω_r 的路径边上的条件概率的乘积. 由于这个原因, 我们可以称鞅概率为路径概率 (path probability).

在实际计算 (6.10.1) 式中的条件概率时, 我们不是看路径而看各独自的块以及它们的直接“后代”, 如图 6 所示. 这就形成了整个模型 $\mathbb{M}$ 的一个子模型.

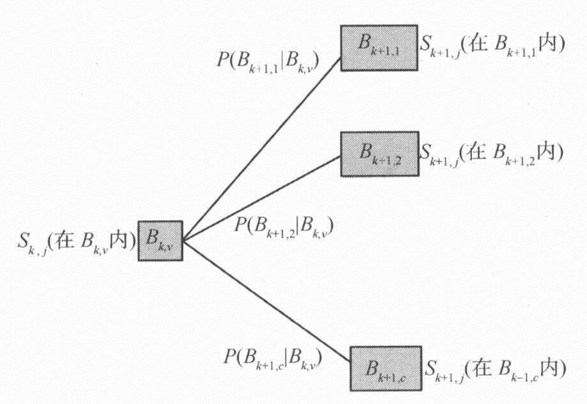


图 6 从 $B_{k,v}$ 处开始的子模型

固定一个块 $B_{k,v}$, 假设从 $B_{k,v}$ 引申出来的块为

$$\mathcal{B} = \{B_{k+1,1}, \cdots, B_{k+1,c}\}.$$

那么对每种资产 a_j, 鞅条件为

$$\overline{S}_{k,j} = \mathcal{E}_\Pi(\overline{S}_{k+1,j}|P_k).$$

随机变量 $\mathcal{E}_\Pi(\overline{S}_{k+1,j}|\mathcal{P}_k)$ 是 $\mathcal{P}_k$ 可测的, 也就是说它在 $\mathcal{P}_k$ 的块上是常数. 因为可有启发性地写成

$$\overline{S}_{k,j}(\text{在 } B_{k,v}\text{内}) = \mathcal{E}_\Pi(\overline{S}_{k+1,j}|\mathcal{P}_k)(\text{在 } B_{k,v}\text{内}),$$

或等价地

$$\overline{S}_{k,j}(\text{在 } B_{k,v}\text{内}) = \mathcal{E}_\Pi(\overline{S}_{k+1,j}|B_{k,v}).$$

由于 $\overline{S}_{k+1,j}$ 在 $\mathcal{P}_{k+1}$ 的块上是常数, 故有

$$\begin{aligned}\overline{S}_{k,j}(\text{在 } B_{k,v}\text{内}) &= \mathcal{E}_\Pi(\overline{S}_{k+1,j}|B_{k,v}) \\ &= \frac{1}{\mathbb{P}_\Pi(B_{k,v})}\mathcal{E}_\Pi(1_{B_{k,v}}\overline{S}_{k+1,j}) \\ &= \frac{1}{\mathbb{P}_\Pi(B_{k,v})}\sum_{i=1}^{c}[1_{B_{k,v}}\overline{S}_{k+1,j}](\text{在 } B_{k,v}\text{内})\mathbb{P}_\Pi(B_{k+1,i}) \\ &= \sum_{i=1}^{c}\overline{S}_{k+1,j}(\text{在 } B_{k+1,i}\text{内})\mathbb{P}_\Pi(B_{k+1,i}|B_{k,v}).\end{aligned}$$

下面的方程 (对每个 $j = 1, \cdots, m$)

$$\overline{S}_{k,j}(\text{在 } B_{k,v}\text{内}) = \sum_{i=1}^{c}\overline{S}_{k+1,j}(\text{在 } B_{k+1,i}\text{内})\mathbb{P}_\Pi(B_{k+1,i}|B_{k,v}).$$

为计算条件概率提供了方法. 注意到从 $j = 1$ 开始, 有 $\overline{S}_{l,1} = 1$, 所以该情形下的方程为

$$1 = \sum_{i=1}^{c}\mathbb{P}_\Pi(B_{k+1,i}|B_{k,v}).$$

定理 10 令 $\Pi = (\pi_1, \cdots, \pi_m)$ 是 $\mathbb{M}$ 上的一个鞅测度. 对每个 $\omega_r \in \Omega$, 存在唯一的块链包含它,

$$B_{0,1} \supseteq B_{1,i_1} \supseteq \cdots \supseteq B_{T,i_T} = \{\omega_r\}.$$

从而鞅概率 $\pi_r = \mathbb{P}_\Pi(\omega_r)$ 刚好是条件概率的乘积

$$\pi_r = \mathbb{P}_\Pi(\omega_r) = \mathbb{P}_\Pi(B_{1,i_1}|B_{0,1})\mathbb{P}_\Pi(B_{2,i_2}|B_{1,i_1})\cdots\mathbb{P}_\Pi(B_{T,i_T}|B_{T-1,i_{T-1}}).$$

为了计算条件概率 $\mathbb{P}_\Pi(\cdot|B_{k,v})$, 假设从 $B_{k,v}$ 引申出来的块为

$$\mathcal{B}=\{B_{k+1,1},\cdots,B_{k+1,c}\}.$$

那么, 得到一个方程系统 (对每个$j=2,\cdots,m$)

$$\overline{S}_{k,j}(\text{在 } B_{k,v}\text{内})=\sum_{i=1}^{c}\overline{S}_{k+1,j}(\text{在 } B_{k+1,i}\text{内})\mathbb{P}_\Pi(B_{k+1,i}|B_{k,v}) \tag{6.10.2}$$

和

$$\sum_{i=1}^{c}\mathbb{P}_\Pi(B_{k+1,i}|B_{k,v})=1. \tag{6.10.3}$$

下面举例说明如何计算鞅测度.

例 2　图 7 左上半部分给出了例 1 的状态树. 假定无风险利率为 0.

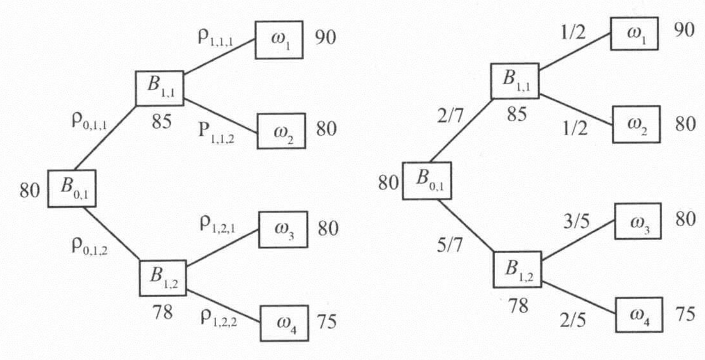


图 7　计算鞅概率

我们可以从倒数第二个划分 $\mathcal{P}_1$ 的每个块开始计算条件概率. 对块 $B_{1,1}$ 来说, 方程 (6.10.2) 和 (6.10.3) 给出

$$\begin{aligned}90\mathbb{P}_\Pi(B_{2,1}|B_{1,1})+80\mathbb{P}_\Pi(B_{2,2}|B_{1,1})&=85,\\ \mathbb{P}_\Pi(B_{2,1}|B_{1,1})+\mathbb{P}_\Pi(B_{2,2}|B_{1,1})&=1.\end{aligned}$$

求解该系统得

$$\mathbb{P}_\Pi(B_{2,1}|B_{1,1})=\mathbb{P}_\Pi(B_{2,2}|B_{1,1})=\frac{1}{2}.$$

如图 7 右边所示. 同样地, 对块 $B_{1,2}$, 可得到

$$\begin{aligned}80\mathbb{P}_\Pi(B_{2,3}|B_{1,2})+75\mathbb{P}_\Pi(B_{2,4}|B_{1,2})&=78,\\ \mathbb{P}_\Pi(B_{2,3}|B_{1,2})+\mathbb{P}_\Pi(B_{2,4}|B_{1,2})&=1.\end{aligned}$$

其解为

$$\mathbb{P}_\Pi(B_{2,3}|B_{1,2})=\frac{3}{5},\quad \mathbb{P}_\Pi(B_{2,4}|B_{1,2})=\frac{2}{5}.$$

最后, 对块 $B_{0,1}$ 有

$$85\mathbb{P}_\Pi(B_{1,1}|B_{0,1})+78\mathbb{P}_\Pi(B_{1,2}|B_{0,1})=80,$$
$$\mathbb{P}_\Pi(B_{1,1}|B_{0,1})+\mathbb{P}_\Pi(B_{1,2}|B_{0,1})=1.$$

其解为

$$\mathbb{P}_\Pi(B_{1,1}|B_{0,1})=\frac{2}{7},\quad \mathbb{P}_\Pi(B_{1,2}|B_{0,1})=\frac{5}{7}.$$

图 7 的右半部分给出了这些条件期望. 现在我们就能通过对每条路径的开始状态到最终状态之间的条件概率相乘来计算鞅测度

$$\mathbb{P}_\Pi(\omega_1)=\frac{2}{7}\cdot\frac{1}{2}=\frac{2}{14},$$
$$\mathbb{P}_\Pi(\omega_2)=\frac{2}{7}\cdot\frac{1}{2}=\frac{2}{14},$$
$$\mathbb{P}_\Pi(\omega_3)=\frac{5}{7}\cdot\frac{3}{5}=\frac{3}{7},$$
$$\mathbb{P}_\Pi(\omega_4)=\frac{5}{7}\cdot\frac{2}{5}=\frac{2}{7}.$$

现在再来考虑未定权益

$$X(\omega_1)=100,\quad X(\omega_2)=90,\quad X(\omega_3)=80,\quad X(\omega_4)=70.$$

一个自融资交易策略 $\Phi=(\Theta_1,\Theta_2)$ 的收益为

$$\mathcal{V}_2(\Theta_2)(\omega_1)=100,$$
$$\mathcal{V}_2(\Theta_2)(\omega_2)=90,$$
$$\mathcal{V}_2(\Theta_2)(\omega_3)=80,$$
$$\mathcal{V}_2(\Theta_2)(\omega_4)=70.$$

这时, 利用定理 7, 该定理告诉我们

$$\overline{\mathcal{V}_0}(\Phi)=\mathcal{E}_\Pi(\overline{\mathcal{V}_T}(\Phi)).$$

因此

$$\overline{\mathcal{V}_0}(\Phi)=100\cdot\frac{2}{14}+90\cdot\frac{2}{14}+80\cdot\frac{3}{7}+70\cdot\frac{2}{7}=\frac{570}{7}\approx 81.43.$$

与例 1 中得到的结果一样.

练　习　6

1. 根据图 4 所给出的状态树, 请求出自融资交易策略 $\Phi = (\Theta_1, \Theta_2)$, 使得该策略复制了下面的未定权益

$$X(\omega_1) = 95, \quad X(\omega_2) = 90, \quad X(\omega_3) = 85, \quad X(\omega_4) = 75.$$

假设无风险利率为 0.

2. 状态树如图 8 所示.

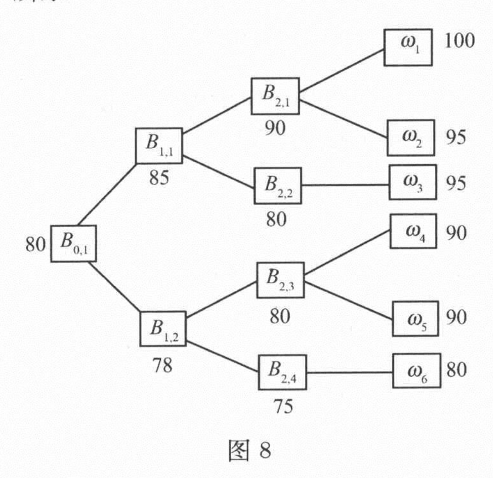


图 8

请求出自融资交易策略 Φ, 使得该策略复制了下面的未定权益

$$(100, 100, 95, 90, 90, 85).$$

假设无风险利率为 0. 提示：可能不止一个答案.

3. 考虑如下的一个赌博. 依次抛 3 枚均匀的硬币. 如果出现 3 次正面, 则玩家赢; 否则庄家 (娱乐场所) 赢. 对于玩家的每 \$0.25 赌注, 庄家需拿出 \$2.00 以使得赌注公平. 现在假想庄家想对冲一个玩家的 \$1 百万赌注 (庄家这时面临 \$8 百万的风险). 因此, 庄家找到一个玩抛硬币赌博的做市商, 采取如下的赌博策略：在抛第一枚硬币之前, 庄家下 \$1 百万赌正面; 在抛第二枚硬币前 (如果存在机会的话), 他下 \$2 百万赌正面; 在抛第三枚硬币前 (如果存在机会的话), 他下 \$4 百万赌正面. 追踪庄家和玩家的组合在赌博过程中的价值. 试证明庄家采取了一种自融资, 完全对冲的复制策略.

4. 证明所有的自融资交易策略形成的集合 $\mathcal{T}$ 在如下的运算规则下是一个向量空间：

• 坐标态加法

$$(\Theta_{1,1},\cdots,\Theta_{1,T})+(\Theta_{2,1},\cdots,\Theta_{2,T})=(\Theta_{1,1}+\Theta_{2,1},\cdots,\Theta_{1,T}+\Theta_{2,T}).$$

• 点乘

$$a(\Theta_1,\cdots,\Theta_T)=(a\Theta_1,\cdots,a\Theta_T).$$

5. 考虑自融资交易策略

$$\Phi=(\Theta_1,\cdots,\Theta_T),$$

其中

$$\Theta_i=(\theta_{i,1},\cdots,\theta_{i,n}).$$

对任何非 0 实数 a, 令

$$\Phi'=(\Theta_1',\cdots,\Theta_T'),$$

其中

$$\Theta_i'=(\theta_{i,1}+a1_\Omega,\cdots,\theta_{i,n}).$$

试证明 Φ' 是自融资的.

6. 证明：由所有的可达到未定权益形成的集合 $\mathcal{M}$ 是向量空间 $RV(\Omega)$ 的一个子空间, 这里 $RV(\Omega)$ 是由 Ω 上所有随机变量形成的空间.

7. 证明定理 4.

8. 考虑一个两资产模型 $\mathbb{M}$：无风险资产和一只股票. 如果无风险利率 r_i 足够大, 是否总存在套利机会? 解释一下你的答案. 该结果能应用到多于一种风险资产的模型中去吗?

9. 考虑如下游戏. 一个盒子里装有 3 枚硬币, 第一枚硬币是均匀的, 第二枚硬币出现正面的概率为 0.55, 第三枚硬币出现正面的概率为 0.45. 画一个状态树, 指出所有可能的结果和它们的概率. 找出路径加权概率分布.

10. 称 $\Phi_1=\Phi_2$ 当且仅当 Φ_1 复制了 Φ_2. 证明由 $\Phi_1=\Phi_2$ 定义的复制关系是自融资交易策略集上的等价关系, 即该关系满足如下条件：

a)(自反性)$\Phi_1=\Phi_1$.

b)(对称性)$\Phi_1=\Phi_2\Rightarrow\Phi_2=\Phi_1$.

c)(传递性)$\Phi_1=\Phi_2$且$\Phi_2=\Phi_3\Rightarrow\Phi_1=\Phi_3$.

11. 证明：如果任何严格正的未定权益是可达到的, 那么市场是完全的.

两资产两状态的单期模型

考虑一个简单的两资产两状态的单期模型 $\mathbb{M}$. 该模型有两种资产 $\mathcal{A}=(a_1,a_2)$, 其中 a_1 是利率为 r 的无风险债券, a_2 是初始价格为 S_0 的标的股票, 其最终价格为 S_T. 该模型仅有两种经济状态 $\Omega=(\omega_1,\omega_2)$. 习惯用股票的初始价格来表示其最终价格. 状态为 ω_1 时, 股价为初始股价的 u 倍, 因此

$$S_T=S_0u.$$

状态为 ω_2 时, 股价为初始股价的 d 倍, 因此

$$S_T=S_0d.$$

假设 $d\leqslant u$. 下面的练习都在该模型中考虑.

12. 证明：$\mathbb{M}$ 是完全的当且仅当 $d<u$.

13. 考虑一个期权, 其收益 X 为

$$X(\omega_1)=f_u,$$
$$X(\omega_2)=f_d.$$

试找出 X 的一个复制组合.

14. 求 X 的初始价格.

15. 令

$$\pi=\frac{\mathrm{e}^{rT}-d}{u-d}.$$

证明该衍生产品的价格为

$$\mathrm{e}^{rT}[\pi f_u+(1-\pi)f_d].$$

从上式出发你如何看待 $(\pi,1-\pi)$?

16. 证明该模型无套利当且仅当 $d<\mathrm{e}^{rT}<u$.

17. 一位日交易员对一特殊股票很感兴趣, 该股票的当前价格为 \$100. 他估计这只股票该天最后的价格为 \$101 或者 \$99. 一个欧式买权以该股票为标的, 执行价为 \$99.5. 该期权的价格为多少? 假设 $r=4\%$.

18. a) 假设某一证券当前售价为 160. 到 T 时, 该证券的售价为 200 或者 140. 假设市场是无套利的且 $r=0$, 请计算以该证券为标的的执行价为 180 的欧式卖权的价格.

b) 假设你非常幸运地能以 20 这个价格获得该卖权. 请描述一个包含该卖权的投资组合, 该组合保证能获得利润.

两资产三状态的单期模型

现在来考虑两资产三状态的单期模型. 假设无风险利率为 0. 假定 $S_{0,2}=25$, 且

$$S_{1,2}(\omega_1)=40,\quad S_{1,2}(\omega_2)=30,\quad S_{1,2}(\omega_3)=20.$$

19. 证明该模型是不完全的.
20. 找出该模型的所有鞅测度.
21. 证明如下测度为鞅测度

$$\begin{aligned}\Pi_1&=\left(\frac{1}{12},\frac{4}{12},\frac{7}{12}\right),\\ \Pi_2&=\left(\frac{1}{6},\frac{1}{6},\frac{4}{6}\right).\end{aligned}$$

22. 找出一个复制交易策略 (投资组合), 用上面给出的两个鞅测度来对执行价为 20 的欧式买权进行定价.

第 7 章　考克斯 – 罗斯 – 鲁宾斯坦 (CRR) 模型

在该章里, 我们将讨论一个特定的离散时间模型, 即考克斯 – 罗斯 – 鲁宾斯坦模型 (Cox-Ross-Rubinstein), 该模型首先由他们在 1979 年提出. 将考克斯 – 罗斯 – 鲁宾斯坦模型简写为 CRR 模型. CRR 模型在著作中也被称为二叉树模型, 通过后面的学习, 我们会发现这种称呼的原因是显然的.

在后面一章, 我们将用该模型来推导著名的布莱克 – 舒尔斯期权定价模型.

7.1　模　　型

时间

CRR 模型是一个离散时间模型, 这样它有有限个时间点

$$t_0 < t_1 < \cdots < t_T.$$

而且, 时间区间 $[t_i, t_{i-1}]$ 的长度相等, 都为 Δt, 即

$$t_i - t_{i-1} = \Delta t.$$

因此, 模型的整个时间长度为

$$L = t_T - t_0 = T\Delta t.$$

资产

CRR 模型中仅有两种资产：无风险资产 a_1 和风险资产 a_2.

模型的状态

图 1 给出了 CRR 模型的状态树的一部分.

CRR 模型假定在每个时间段 $[t_i, t_{i+1}]$ 中, 经济状态只以两种方式中的一种来改变：要么上涨, 要么下跌. 也假定改变方向与过去的改变方向独立.

如果记上涨为 U, 下跌为 D, 那么经济的最后状态就是由 U 和 D 组成的长度为 T 的字符串. 记由 U 和 D 组成的长度为 k 的字符串集合为 $\{U, D\}^k$. 例如

$$\{U, D\} = \{UU, UD, DU, DD\}.$$

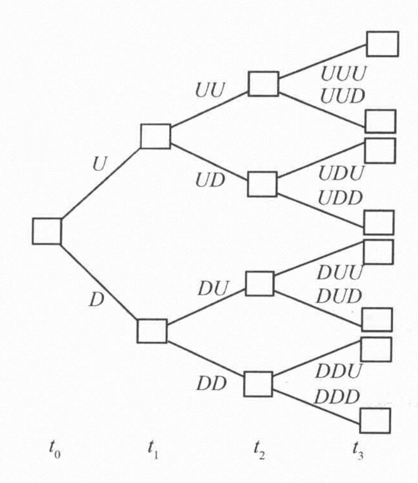


图 1　CRR 模型的状态树

因此, 最后的状态空间为

$$\Omega = \{U, D\}^T.$$

注意 $\{U, D\}^k$ 的长度为 2^k, 特别地, Ω 的长度为 2^T.

因为我们将会经常处理 U 和 D 形成的字符串, 所以先来确定一些记号. 对任意 $\omega \in \{U, D\}^T$, 记 ω 中长度为 i 的前缀为 $[\omega]_i$. 因此, 如果 $\omega = e_1 \cdots e_T$, 那么对任何 $i \leqslant T$,

$$[\omega]_i = e_1 \cdots e_i.$$

令

$$N_U(\omega) = \omega \text{ 中}U\text{的个数},$$
$$N_D(\omega) = \omega \text{ 中}D\text{的个数}.$$

定义模型的中间状态如下. 对 $\{U, D\}^k$ 中的每个字符串, 存在一个时刻 t_k 的中间状态. 特别地, 对 $\delta = e_1 \cdots e_k \in \{U, D\}^k$, 中间状态 $B_\delta \in \mathcal{P}_k$ 为有前缀 δ 的所有最终状态

$$B_\delta = \{\omega \in \Omega | [\omega]_i = \delta\}.$$

因此, $\mathcal{P}_k$ 刚好有 2^k 个块 (中间状态).

例如, 如果 $T = 4$, 那么 $\mathcal{P}_1$ 由两个中间状态组成

$$B_U = \{UUUU, UUUD, UUDU, UUDD, UDUU, UDUD, UDDU, UDDD\},$$
$$B_D = \{DUUU, DUUD, DUDU, DUDD, DDUU, DDUD, DDDU, DDDD\}.$$

$\mathcal{P}_2$ 由四个中间状态组成

$$
\begin{aligned}
B_{UU} &= \{UUUU, UUUD, UUDU, UUDD\},\\
B_{UD} &= \{UDUU, UDUD, UDDU, UDDD\},\\
B_{DU} &= \{DUUU, DUUD, DUDU, DUDD\},\\
B_{DD} &= \{DDUU, DDUD, DDDU, DDDD\}.
\end{aligned}
$$

在时刻 t_0, 只有一个 (初始) 状态 $B_\varepsilon = \Omega$. 这对应着空字符串 ε, 它为所有字符串的前缀.

非常清楚的是, $\mathcal{P}_k$ 中的每个块 $B_{e_1\cdots e_k}$ 刚好为下一个划分 $\mathcal{P}_{k+1}$ 引出两个块, 即

$$B_{e_1\cdots e_k U}, B_{e_1\cdots e_k D}.$$

换一种方式说, 状态树的每个结点刚好引出两个节点.

客观概率

我们也需要考虑任何时期经济上升趋势的客观概率. 记该概率为 p. 需强调的是, 客观概率是经济上的估计而不是数学意义上的.

价格函数

为了使记号简单化, 记 t_k 时无风险资产的价格为 B_k, 风险资产的价格为 S_k, 具体可以将风险资产想象为一只股票.

CRR 模型指定股价由一对实数 u 和 d 决定, u 和 d 满足

$$0 < d < u.$$

如果经济在 $[t_i, t_{i+1}]$ 内上升, 那么股价从 S_k 上涨到 $S_k u$; 如果经济下滑, 那么股价也从 S_k 下跌到 $S_k d$. 注意 u 和 d 都是常数, 即它们不依赖时间.

从而 t_k 时的股票价格函数 S_k 由如下给出

$$S_k(\omega) = S_0 u^{N_U([\omega]_k)} d^{N_D([\omega]_k)}.$$

对任何最终状态 $\omega \in \{U, D\}^T$. 特别地, 最后的价格为

$$S_T(\omega) = S_0 u^{N_U(\omega)} d^{N_D(\omega)}.$$

S_k 是 $\mathcal{P}_k$ 可测的这个事实可由如下事实反映出来. 价值 $S_k(\omega)$ 只依赖 ω 的前缀 $[\omega]_k$, 故仅依赖最初到 t_k 这段时间内所发生的事情. 也注意到股票在 t_k 时的价格仅

依赖这时状态中 U 和 D 的个数, 而与 U 和 D 的顺序无关. 这就是 CRR 模型具有而一般离散时间模型所不具有的重要特征 (虽然也可能不太现实).

注意股票价格函数也满足一个循环关系

$$S_k(\omega) = S_{k-1}u^{E_k(\omega)}d^{1-E_k(\omega)}.$$

无风险资产的价格总可由无风险利率给出. 在 CRR 模型中, 假设在模型的整个过程中无风险利率 r 为常数. 因此, 对任何最终状态 ω 来说, 无风险资产在 t_k 时的价格为

$$\mathrm{e}^{r(t_k-t_0)}.$$

当然, 单位必须相配. 比如, 如果 r 是年利率, 那么时间 t_k 也必须用年来衡量.

7.2 CRR 模型中的鞅测度

假设 Π 是 CRR 模型 $\mathbb{M}$ 中的一个鞅测度. 第 6 章中的定理 10 告诉我们如何计算条件概率, 这些条件概率被用来计算 Π.

考虑 $\mathcal{P}_k$ 中的一个块

$$B_\delta = \{\omega \in \Omega | [\omega]_i = \delta\},$$

B_δ 所包含 $\mathcal{P}_{k+1}$ 中的块为

$$B_{\delta U} = \{\omega \in \Omega | [\omega]_{i+1} = \delta U\}$$

和

$$B_{\delta D} = \{\omega \in \Omega | [\omega]_{i+1} = \delta D\}.$$

图 2 给出了块 B_δ 和它的直接后代.

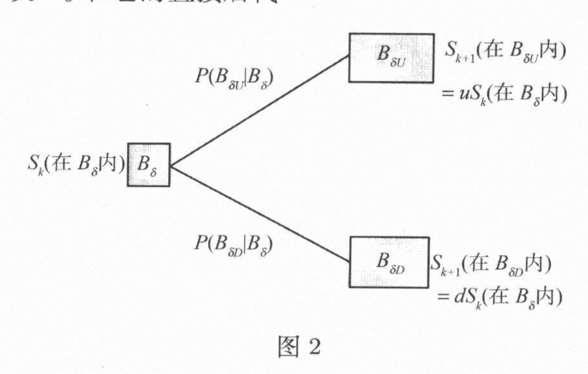


图 2

记条件概率为

$$\rho_{\delta,U} = \mathbb{P}_\Pi(B_{\delta U}|B_\delta),$$
$$1-\rho_{\delta,U} = \mathbb{P}_\Pi(B_{\delta D}|B_\delta).$$

CRR 模型规定

$$S_{k+1}(\text{在 } B_{\delta U}\text{内}) = uS_k(\text{在 } B_\delta\text{内}),$$
$$S_{k+1}(\text{在 } B_{\delta D}\text{内}) = \mathrm{d}S_k(\text{在 } B_\delta\text{内}).$$

或者用折现形式表示为 (两边乘以 $\mathrm{e}^{-(k+1)r\Delta t}$)

$$\overline{S}_{k+1}(\text{在 } B_{\delta U}\text{内}) = \mathrm{e}^{-r\Delta t}u\overline{S}_k(\text{在 } B_\delta\text{内}),$$
$$\overline{S}_{k+1}(\text{在 } B_{\delta D}\text{内}) = \mathrm{e}^{-r\Delta t}d\overline{S}_k(\text{在 } B_\delta \text{ 内}).$$

由第 6 章的定理 10 得

$$\overline{S}_k(\text{在 } B_\delta\text{内}) = \mathrm{e}^{-r\Delta t}u\overline{S}_k(\text{在 } B_\delta\text{内})[u\rho_{\delta,U} + d(1-\rho_{\delta,U})],$$

或者

$$\mathrm{e}^{r\Delta t} = u\rho_{\delta,U} + d(1-\rho_{\delta,U}) = (u-d)\rho_{\delta,U} + d.$$

由于 $\rho_{\delta,U}$ 与 δ 独立, 故可以写成

$$\pi_U = \rho_{\delta,U}.$$

有

$$\pi_U = \frac{\mathrm{e}^{r\Delta t} - d}{u-d}$$

和

$$1-\pi_U = \frac{u-\mathrm{e}^{r\Delta t}}{u-d}.$$

现在, 序对 $(\pi_U, 1-\pi_U)$ 是强正的概率分布当且仅当 $0 < \pi_U < 1$. 在这种情形下, 条件概率仅依赖 u, d 和 r, 且是唯一的. 这就暗示着鞅测度是唯一的, 因此模型是完全的.

条件 $0 < \pi_U < 1$ 等价于

$$0 < \mathrm{e}^{r\Delta t} - d < u - d,$$

或者

$$d < \mathrm{e}^{r\Delta t} < u.$$

假设就是这种情况, 得出的唯一鞅测度 Π 由如下给出：对任意 $\omega \in \{U, D\}^T$,

$$\mathbb{P}_\Pi(\omega) = \pi_U^{N_U(\omega)}(1-\pi_U)^{T-N_U(\omega)}.$$

现在我们有一个非常完美的定理来描述 CRR 模型中的鞅测度.

定理 1 CRR 模型 $\mathbb{M}$ 是完全的和无套利的当且仅当

$$d < \mathrm{e}^{r\Delta t} < u.$$

在该情形下, $\mathbb{M}$ 上的唯一鞅测度 Π 定义如下：对任意 $\omega \in \{U, D\}^T$,

$$\mathbb{P}_\Pi(\omega) = \pi_U^{N_U(\omega)}(1-\pi_U)^{T-N_U(\omega)},$$

其中

$$\pi_U = \frac{\mathrm{e}^{r\Delta t} - d}{u - d}.$$

7.3 在 CRR 模型中的定价

假定 $\mathbb{M}$ 是完全且无套利的 CRR 模型. 从而可用复制策略来对未定权益进行定价. 特别地, 如果 X 是一个未定权益, 那么存在一个交易策略 Φ 复制 X, 故 X 的价格为

$$\begin{aligned}
\mathcal{I}(X) &= \mathcal{V}_0(\Phi) \\
&= \mathrm{e}^{-rL}\mathcal{E}_\Pi(\mathcal{V}_T(\Phi)) \\
&= \mathrm{e}^{-rL}\mathcal{E}_\Pi(X) \\
&= \mathrm{e}^{-rL}\sum_{\omega\in\Omega} X(\omega)\mathbb{P}_\Pi(\omega) \\
&= \mathrm{e}^{-rL}\sum_{\omega\in\Omega} X(\omega)\pi_U^{N_U(\omega)}(1-\pi_U)^{T-N_U(\omega)}.
\end{aligned}$$

对一个股票期权来说, 比如买权, 它的最终损益为

$$\mathcal{V}_T = \max\{S_T - K, 0\} = (S_T - K)^+,$$

这里 x^+ 表示 $\max\{x, 0\}$.

现在, 在具有相同数目 U 的所有最终状态下, 股价 S_T 是一样的, 事实上

$$S_T(\omega) = S_0 u^{N_U(\omega)} d^{T-N_U(\omega)}.$$

因此, 可以基于 ω 中 U 的个数来重编上面最终和的各项. 因为长度为 T 且刚好有 k 个 U 的序列的个数为

$$C_T^n = \frac{n!}{k!(n-k)!}.$$

所以有

$$\mathrm{e}^{-rL}\sum_{\omega\in\Omega} X(\omega)\pi_U^{N_U(\omega)}(1-\pi_U)^{T-N_U(\omega)}$$

$$=\mathrm{e}^{-rL}\sum_{\omega\in\Omega}(S_0u^{N_U(\omega)}d^{T-N_U(\omega)}-K)^+\pi_U^{N_U(\omega)}(1-\pi_U)^{T-N_U(\omega)}$$

$$=\mathrm{e}^{-rL}\sum_{k=0}^{T}C_T^k(S_0u^kd^{T-k}-K)^+\pi_U^k(1-\pi_U)^{T-k}.$$

由于该公式非常重要, 故我们将它放到一个定理中去.

定理 2　假设 $\mathbb{M}$ 是一个完全且无套利的 CRR 模型. 那么在模型的最后时刻到期且执行价为 K 的欧式买权的初始价值为

$$\mathcal{I}(\text{买权})=\mathrm{e}^{-rL}\sum_{k=0}^{T}C_T^k(S_0u^kd^{T-k}-K)^+\pi_U^k(1-\pi_U)^{T-k}.$$

欧式卖权的初始价值为

$$\mathcal{I}(\text{卖权})=\mathrm{e}^{-rL}\sum_{k=0}^{T}C_T^k(K-S_0u^kd^{T-k})^+\pi_U^k(1-\pi_U)^{T-k}.$$

例 1　某股票当前价格为 100. 感觉在接下来的两个月中的每一个月里, 股价会上涨 1% 或下跌 1%. 假设无风险利率为 1%, 对不同的执行价 $K=102, K=101, K=100, K=99, K=98, K=97$, 请计算欧式卖权的价格.

解　鞅概率为

$$\pi_U=\frac{\mathrm{e}^{r\Delta t}-d}{u-d}=\frac{\mathrm{e}^{(0.01)(1/12)}-0.99}{0.02}\approx 0.54.$$

故

$$\mathcal{I}=\mathrm{e}^{-0.01/6}\sum_{k=0}^{2}C_2^k(100(1.01)^k(0.99)^{2-k}-K)^+(0.54)^k(0.46^{2-k})$$

$$=0.9983[0.2116(98.01-K)^+ +0.4968(99.99-K)^+ +0.2916(102.01-K)^+].$$

因此, 通过计算得出

K	102	101	100	99	98	97
S_0	0.0029	0.2959	0.5888	1.3725	2.1632	3.1615

7.4　从另一角度看 CRR 模型与随机游走

我们从一个稍微不同的观点来看 CRR 模型. 在每个时间段 $[t_i, t_{i+1}]$(区间长度为 Δt) 内, 股价要么上涨要么下跌. 因此, 单个的价格运动可看作一列独立的伯努利随机变量 E_i, 其分布为

$$\mathbb{P}(E_i = u) = p,$$
$$\mathbb{P}(E_i = \mathrm{d}) = 1 - p.$$

也就是说,

$$E_i = \begin{cases} u, & \text{以概率 } p, \\ d, & \text{以概率 } q = 1 - p, \end{cases}$$

这里 p 为经济上升的客观概率. 因此最后时刻 t_T 的股价为

$$S_T = S_0 E_1 \cdots E_T = S_0 \mathrm{e}^{\sum \log(E_i)} = S_0 \mathrm{e}^{H_T},$$

这里

$$H_T = \log\left(\frac{S_T}{S_0}\right) = \sum_{i=1}^{T} \log(E_i)$$

为股价的对数增长率. 接下来定义常数 μ 和 σ 如下

$$\mu = \frac{1}{\Delta t}\mathcal{E}(\log E_i) = \frac{1}{\Delta t}(p \log u + q \log d),$$
$$\sigma^2 = \frac{1}{\Delta t}\mathrm{var}(\log E_i) = \frac{1}{\Delta t} pq(\log u - \log d)^2.$$

后面会讨论这些常数的重要性. 现在, 因为有

$$\mathcal{E}(\log E_i) = \mu \Delta t$$

和

$$\mathrm{var}(\log E_i) = \sigma^2 \Delta t.$$

对随机变量 $\log E_i$ 进行标准化, 得到 (因为 $\sigma \neq 0$)

$$X_i = \frac{\log E_i - \mu \Delta t}{\sigma \sqrt{\Delta t}}.$$

现在将 $\log E_i$ 写成

$$\log E_i = \mu \Delta t + \sigma \sqrt{\Delta t}\left[\frac{\log E_i - \mu \Delta t}{\sigma \sqrt{\Delta t}}\right] = \mu \Delta t + \sigma \sqrt{\Delta t} X_i,$$

其中随机变量

$$X_i = \frac{\log E_i - \mu \Delta t}{\sigma \sqrt{\Delta t}}$$

是独立的伯努利随机变量, 有

$$X_i = \begin{cases} \dfrac{q}{\sqrt{pq}}, & \text{以概率 } p, \\ \dfrac{-p}{\sqrt{pq}}, & \text{以概率 } q. \end{cases}$$

因此

$$\mathcal{E}(X_i) = 0,$$

$$\operatorname{var}(X_i) = 1.$$

现在有

$$\begin{aligned} H_T &= \sum_{i=1}^{T} \log(E_i) \\ &= \sum_{i=1}^{T} [\mu\Delta t + \sigma\sqrt{\Delta t} X_i] \\ &= \mu L + \sigma\sqrt{\Delta t} \sum_{i=1}^{T} X_i, \end{aligned}$$

即

$$H_T = \mu L + \sigma\sqrt{\Delta t} \sum_{i=1}^{T} X_i.$$

该公式表明对数增长率为确定部分 μL 和随机部分的和, 确定部分为模型的整个时间长度 L 乘以一个常数, 随机部分为

$$\sigma\sqrt{\Delta t} \sum_{i=1}^{T} X_i,$$

这是独立伯努利随机变量的和乘以一个常数. 每项 X_i 描述了股价在模型的一个子区间内的运动. 最后, 股价本身可由如下给出

$$S_T = S_0 \mathrm{e}^{H_T} S_0 \mathrm{e}^{\mu L + \sigma\sqrt{\Delta t} \sum_{i=1}^{T} X_i}.$$

称常数 μ 为股价的漂移系数 (drift), σ 为股价的波动率 (volatility). 这两个术语被用来解释股价瞬间变化情况.

注意, 有些作者称如下表达式为回报收益

$$s = \frac{1}{T} \log\left(\frac{S_T}{S_0}\right) = \frac{1}{T} H_T.$$

这是因为上面的等式等价于

$$S_T = S_0 \mathrm{e}^{sT}.$$

这表明股价以连续复合利率 s 增长. 因此, s 是收益率.

随机游走

用来描述股价在每个子区间内的表现行为的序列 (X_i) 是随机游走的一个例子. 为了更好地理解随机游走, 假设一只跳蚤只能在一条直线上跳, 比如说 X 轴. 在 $t=0$ 时, 跳蚤从点 $x=0$ 出发, 在每一个时间间隔 (长为 Δt) 内随机向右或向左跳, 每次向右跳的距离为 a, 向左跳的距离为 b. 进一步假定向右跳的概率为 p. 如图 3 所示.

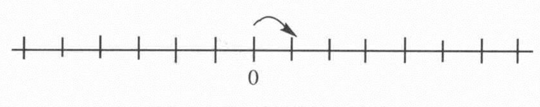


图 3　跳蚤的随机游走

序列 (X_i) 中的每个随机变量 X_i 描述了跳蚤行为过程中的一简单步伐, 下面的部分和表示了跳蚤在时刻 t_k 所处的位置

$$U_k = \sum_{i=1}^{k} X_i.$$

图 4 给出了由计算机生成的一对随机游走图, 这里 $p=q=1/2$ 且 $a=b$(叫做对称随机游走). 为了能清楚地分析路径, 习惯性在平面上把跳蚤的每个位置用一个点标识, 这里 X 轴表示时间, Y 轴表示位置.

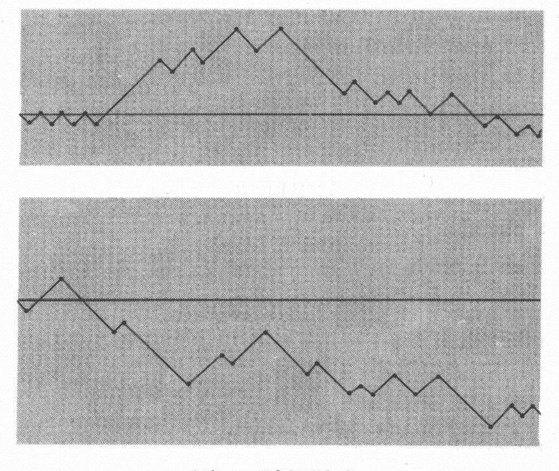


图 4　随机游走

有许多有关随机游走现象的简单描述, 比如, 喝醉了的酒鬼沿着一条街随机走, 或者赌徒参与一个结果是随机的赌博或者股票价格的变动. 确实, 在随机游走方面有完整的论著.

对一个随机游走的行为, 我们可以问许多问题. 例如, 给定整数 a, b, 其中 $a < 0 < b$, 一只跳蚤最后一定会到达两个"边界点"中的一个吗? 还是它在边界点之间来回游荡而永远不会到达边界点?

因为前面问题的答案是跳蚤最终一定会到达一个边界点, 所以我们可能问到达每个边界点的概率是多少, 到达一个边界点的期望时间是多少. 我们也可以问, 跳蚤是否一定会在未来某个时刻回到原点.

不管如何, 该书不是一本专门讲随机游走的书, 因此重新回到身边的问题, 即

$$H_T = \mu L + \sigma\sqrt{\Delta t}\sum_{i=1}^{T} X_i,$$

确定项 μL 为模型的整个周期长度乘以一个常数, 它说明了股价的一个稳定变化 (如果 $\mu \neq 0$), 类似于利率为 μ 的无风险资产的表现. 随机项为随机游走项之和乘以一个常数.

下面来总结一下刚刚学习过的有关 CRR 模型的内容. 在后面的一章里, 我们将用该模型来推导著名的布莱克 – 舒尔斯期权定价公式.

定理 3 考虑一个 CRR 模型, 股价上涨的概率为 p, 下跌的概率为 $q = 1 - p$, 模型整个周期长为 L, 每个时间间隔长为 Δt, 那么股价的表达式如下

$$S_T = S_0 \mathrm{e}^{\mu L + \sigma\sqrt{\Delta t}\sum_{i=1}^{T} X_i},$$

其中漂移系数和波动率的定义如下

$$\mu = \frac{1}{\Delta t}(p\log u + q\log d),$$
$$\sigma^2 = \frac{1}{\sqrt{\Delta t}} pq(\log u - \log d)^2.$$

股票运动的随机游走部分为

$$Y_T = \sum_{i=1}^{T} X_i,$$

这里随机变量 X_i 相互独立, 且

$$X_i = \begin{cases} \dfrac{q}{\sqrt{pq}}, & \text{以概率} p, \\ \dfrac{-p}{\sqrt{pq}}, & \text{以概率} q. \end{cases}$$

练 习 7

1. 某股票的当前卖价为 50, 感觉在接下来的两个月中的每一个月里, 股价会上涨 10% 或下跌 10%. 假设无风险利率为 1%, 请计算具有不同执行价的欧式买权的价格. K 的取值如下：

a) 52; b) 51; c) 50; d) 49; e) 48; f) 47.

对有相同执行价和到期日的欧式卖权, 情况又如何呢?

2. 某股票的当前卖价为 50, 感觉在接下来的两个月中的每一个月里, 股价会上涨 5% 或下跌 10%. 假设无风险利率为 1%, 请计算具有不同执行价的欧式买权的价格. K 的取值如下：

a) 11; b) 10; c) 9; d) 8.

对有相同执行价和到期日的欧式卖权, 情况又如何呢?

3. 参看例 1, 请解释为什么所有状态都有损失, 除了第一个状态, 也就是说存在损失的概率是 3/4.

4. 证明 $\{U, D\}^k$ 的长度为 2^k. 提示：利用数学归纳法或者计数基本原理 (也称为乘法规则).

5. 证明：

$$X_i = \begin{cases} \dfrac{q}{\sqrt{pq}}, & \text{以概率}p, \\ \dfrac{-p}{\sqrt{pq}}, & \text{以概率}q. \end{cases}$$

6. 证明：一个参数为 $p = 1/2$ 的伯努利随机变量 X 的两个取值为 $\mathcal{E}(X) = \pm\sqrt{\text{var}(X)}$.

7. 假设未定权益 X 只依赖最终状态中 U 的个数, 则称这样的未定权益为路径独立的未定权益 (path-independent alternative). 特别地, 如果 $\mathcal{P}$ 是 Ω 的一个划分, 它所包含的块为 Ω 的子集 G_k, G_k 刚好含有 k 个 U, 即

$$G_k = \{\omega \in \Omega | N_U(\omega) = k\}.$$

那么 X 是路径独立的当且仅当存在常数 X_k, 使得对 $k = 0, \cdots, T$, 有

$$X_k = X, \quad \forall \omega \in G_k.$$

a) 证明：

$$|G_k| = C_T^k.$$

b) 证明在鞅测度下, 任何 $\omega \in G_k$ 的概率为

$$\pi_U^{N_U(\omega)}(1-\pi_U)^{T-N_U(\omega)} = \pi_U^k(1-\pi_U)^{T-k}.$$

c) 证明事件 G_k 的概率为

$$\mathbb{P}_\Pi(G_k) = C_T^k \pi_U^k (1-\pi_U)^{T-k}.$$

d) 证明：如果 X 是路径独立的未定权益, 那么

$$\mathcal{I}(X) = \mathrm{e}^{-rL} \sum_{k=0}^{T} X_k C_T^k \pi_U^k (1-\pi_U)^{T-k}.$$

8. 根据 $T=2$ 情况下的 CRR 模型编一个计算机程序或者一个 Excel 表格, 可用来计算欧式买权的价格.

9. 验证：

$$\mathcal{E}_p(\log E_i) = p\log u + q\log d,$$
$$\mathrm{var}_p(\log E_i) = pq(\log u - \log d)^2.$$

10. 在一般的离散时间模型中, 知道了一给定时间的经济状态, 这就暗示着知道了该时间的资产价格. 为什么? 反过来成立吗? 如果在时刻 t_k 我们知道了所有以前的状态和资产价格, 又会怎样呢? 论证你的答案. 如果是在 CRR 模型中考虑, 又会怎么样呢?

第 8 章 概率论 3：连续概率

在这一章里, 我们将讨论一般情形下的概率论的一些概念, 这里对样本空间没有限制, 如有限或离散等. 这将作为讨论布莱克 – 舒尔斯衍生品定价模型的一些准备知识.

因为这不是一本有关概率论的书, 对概率论的详细讨论会引导我们偏离自己的主要目标, 所以我们在讨论中就比较“粗糙”. 如果想更详细地了解概率论, 可参考本书最后列出的一些参考文献.

8.1 一般的概率空间

先回顾一下有限概率空间的定义.

定义 1 有限概率空间是一个二元对 $(\Omega, \mathbb{P})$, 其中 Ω 为有限非空集合, 称为样本空间; $\mathbb{P}$ 是定义在 Ω 的所有子集上的实值函数, 称为 Ω 上的概率测度. 函数 $\mathbb{P}$ 必须满足如下性质:

1) (值域) 对所有的 $A \subseteq \Omega$, 有

$$0 \leqslant \mathbb{P}(A) \leqslant 1.$$

2) (全集的概率)

$$\mathbb{P}(A) = 1.$$

3) (可加性) 如果 A 和 B 不相交, 则有

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

在上下文中, Ω 的子集被称作事件.

我们也看到 $\mathbb{P}$ 的可加性等价于有限可加性, 也就是说, 如果

$$A_1, A_2, \cdots, A_n$$

是两两不相交的事件, 那么

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) = \mathbb{P}(A_1) + \cdots + \mathbb{P}(A_n).$$

现在我们想将该定义推广到任意尺度的样本空间, 而尽量保持当前定义的本质. 特别地, 最起码的是 $\mathbb{P}$ 不仅要满足上述三个性质, 而且还要是可数可加的, 即如果

$$A_1, A_2, \cdots$$

是两两不相交的事件, 那么

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i),$$

这里右边的无穷和必须收敛. 我们将尽可能使前面的定义满足可数可加性.

结果表明这可通过折中的办法做到, 也就是说, 不是样本空间的所有子集被考虑为事件. 换句话说, 一般不可能在无穷集 Ω 的所有子集上定义一个可数可加的集合函数. 我们很想给出一个例子来支持该声明, 但是这样的例子涉及太多的数学原理而不适合本书. 因此我们必须要求读者在心里记住这一点.

基于上述事实, 我们考虑这样一个问题: 样本空间的子集的什么类型的群集能作为一个概率测度的事件的群集. 这把我们带入到 σ 代数.

定义 2　设 Ω 是非空集. 称由 Ω 中的子集形成的非空群集 Σ 是一个 σ 代数, 如果

1) $\Omega \in \Sigma$.

2) Σ 对可数并运算封闭, 也就是说, 如果 $A_1, A_2, \cdots$ 是 Σ 中元素形成的序列, 那么

$$\bigcup_{i=1}^{\infty} A_i \in \Sigma.$$

3) Σ 对取余运算封闭, 也就是说, 如果 $A \in \Sigma$, 那么 $A^c \in \Sigma$.

注意到 $\varnothing \in \Omega^c \in \Sigma$. 同样, DeMorgan 定律表明 Σ 对可数交运算也是封闭的 (我们将细节留作练习).

定义 3　可测空间是一个序对 (Ω, Σ), 其中 Ω 是非空集, Σ 是 Ω 上的 σ 代数.

现在可以定义一般的概率空间了.

定义 4　一个概率空间是一个三元组 $(\Omega, \Sigma, \mathbb{P})$, 其中 Ω 是一个非空集, 称为样本空间, Σ 是 Ω 上的 σ 代数, 它的元素称为事件, $\mathbb{P}$ 是定义在 Σ 上的实值函数, 称为概率测度. 函数 $\mathbb{P}$ 必须满足下面性质:

1) (值域) 对任何 $A \in \Omega$, 有

$$0 \leqslant \mathbb{P}(A) \leqslant 1.$$

2) (全集的概率)

$$\mathbb{P}(\Omega) = 1.$$

3) (可数可加性)

$$A_1, A_2, \cdots$$

是两两互不相交的事件, 则

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

下一定理给出了概率测度一个非常有用的性质. 称一个事件序列是单调降的, 如果该序列满足

$$A_1 \supseteq A_2 \supseteq \cdots.$$

同样地, 称一个事件序列是单调增的, 如果该序列满足

$$A_1 \subseteq A_2 \subseteq \cdots.$$

定理 1 在下述意义下, 概率测度是单调连续的:

1) 如果 $A_1 \supseteq A_2 \supseteq \cdots$ 是单调降序列, 那么

$$\lim_{i\to\infty} \mathbb{P}(A_i) = \mathbb{P}\left(\bigcap_{i=1}^{\infty} A_i\right).$$

2) 如果 $A_1 \subseteq A_2 \subseteq \cdots$ 是单调增序列, 那么

$$\lim_{i\to\infty} \mathbb{P}(A_i) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right).$$

证明 对 1) 部分, 假设 $A_1 \supseteq A_2 \supseteq \cdots$. 那么概率 $\mathbb{P}(A_i)$ 序列是非增的实数序列, 有下界 0. 由初等实分析的一个定理知, 这样的序列一定收敛, 因此问题中的极限存在.

关于收敛, 令 $A = \bigcap_{i=1}^{\infty} A_i$. 首先考虑事件

$$A_1 \backslash A_2, \quad A_2 \backslash A_3, \cdots.$$

这些事件互不相交, 因为如果 $i < j$, 那么 $i+1 \leqslant j$, 所以 $A_j \subseteq A_{i+1}$. 从而, 如果

$$a \in (A_i \backslash A_{i+1}) \cap (A_j \backslash A_{j+1}).$$

那么 a 会在 A_j 中而不会在 A_{i+1} 中. 同样, 这样的每个事件与交集 A 也是不相交的. 因此, A_1 可以写成不相交的集合的并

$$A_1 = \left(\bigcup_{i=1}^{\infty} (A_i \backslash A_{i+1})\right) \cup A.$$

如果 $a \in A_1$ 而 $a \notin A$, 那么令 $i+1$ 是第一个下标, 使得 $a \notin A_{i+1}$. 这就得到 $a \in A_i \backslash A_{i+1}$.

现在, 我们运用可数可加性得到

$$\begin{aligned}
\mathbb{P}(A_1) &= \mathbb{P}\left(\bigcup_{i=1}^{\infty}(A_i \backslash A_{i+1})\right) + \mathbb{P}(A) \\
&= \sum_{i=1}^{\infty} \mathbb{P}(A_i \backslash A_{i+1}) + \mathbb{P}(A) \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{P}(A_i \backslash A_{i+1}) + \mathbb{P}(A) \\
&= \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{i=1}^{n}(A_i \backslash A_{i+1})\right) + \mathbb{P}(A) \\
&= \lim_{n\to\infty} \mathbb{P}(A_1 \backslash A_{n+1}) + \mathbb{P}(A) \\
&= \lim_{n\to\infty} [\mathbb{P}(A_1) - \mathbb{P}(A_{n+1})] + \mathbb{P}(A) \\
&= \mathbb{P}(A_1) - \lim_{n\to\infty} \mathbb{P}(A_{n+1}) + \mathbb{P}(A).
\end{aligned}$$

因此

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} \mathbb{P}(A_{n+1}) = \mathbb{P}(A).$$

这就是要证的. 关于 2) 的证明, 留作练习.

8.2 $\mathbb{R}$ 上的概率测度

从理论和应用两方面来看, 最重要的样本空间就是实直线 $\mathbb{R}$. 事实上, 在本书中要考虑的唯一的无限概率空间就是 $\mathbb{R}$.

$\mathbb{R}$ 上最重要的 σ 代数就是 Borel σ 代数 $\mathcal{B}$. Borel σ 代数的正式定义是简单的, 但领会它不一定那么简单.

定义 5 $\mathbb{R}$ 上的 Borel σ 代数 $\mathcal{B}$ 是 $\mathbb{R}$ 上包含所有开区间 (a, b) 的最小 σ 代数, 其中 $a, b \in \mathbb{R}$.

我们来审查该定义. 首先, 必须指出存在一个这样的 σ 代数. 毕竟, 我们用了这样的词汇 “最小的集合 ……”, 这并不意味着一定存在这样的集合.

证明具有某种性质的最小集合存在, 其一般步骤就是指出两件事: 第一, 至少存在一个集合具有所要求的性质; 第二, 具有该性质的任何集合的交也具有该性质. 这样具有所要求性质的所有集合的交存在, 且是具有所要求性质的最小集合.

关于我们要考虑的问题, 容易知道至少存在 $\mathbb{R}$ 上的一个 σ 代数包含所有的开区间：包含 $\mathbb{R}$ 所有子集的群集. 另外, 不难发现 σ 代数的交也是 σ 代数. 我们将该细节留给读者. 因此, Borel σ 代数确实存在, 且是所有包含开区间的 σ 代数的交.

注意, 在确定 Borel σ 代数存在的同时, 所有包含开区间的 σ 代数的交这个描述不太实用. 从实用性角度考虑, 考虑 $\mathcal{B}$ 中元素 (Borel 集) 的一些例子更有用.

定理 2 1) 所有的开区间, 闭区间和半开闭区间都是 Borel 集.

2) 所有半直线 $(-\infty,b],(-\infty,b),[a,\infty),(a,\infty)$ 都是 Borel 集.

3) 所有的开集和闭集都是 Borel 集.

证明 我们粗略地看看证明. 对 1), 为了指出半开闭区间 $(a,b]$ 是一个 Borel 集, 观察到

$$(a,b]=\bigcap_{n=1}^{\infty}\left(a,b+\frac{1}{n}\right).$$

因此 $(a,b]$ 是开区间的可数并, 故也在 $\mathcal{B}$ 中.

对 3), 简要地讨论 $\mathbb{R}$ 中的开集. 称 $\mathbb{R}$ 中的子集 A 是开集, 如果对任一 $x\in A$, 存在一个开区间 (a,b), 使得

$$x\in(a,b)\subseteq A.$$

称一个集合是闭集, 如果它的补集是开集. 设 A 是 $\mathbb{R}$ 中的一个开集. 那么 A 是所有被 A 包含的开区间的并. 事实上, A 是 A 中所有最大的开区间的并. 称 A 中的一个开区间 I 在 A 中是最大的, 如果没有其他包含 I 的开区间也是 A 的真子集.

现在, 我们断定任何两个最大的开区间是不相交的, 且至多存在可数个最大的开区间. 关于前者, A 的任何两个不同的最大开区间一定不相交, 否则它们的并将会是 A 的一个严格更大的开区间. 结果 A 的每个最大开区间包含一个不同的有理数, 又因为有理数是可数的, 所以在 A 中至多存在可数个最大的开区间.

因此, A 是至多可数个开区间的并, 从而它是一个 Borel 集.

最后, 既然所有的开集都是 Borel 集, 又因为闭集是一个开集的补, 所以所有的闭集也是 Borel 集.

首先, 对 Borel 集想得越多就越感到 $\mathbb{R}$ 的所有子集都是 Borel 集. 尽管如此, 还是存在不是 Borel 集的情形. 但是, 大部分 "非病态的" 集合确实是 Borel 集. 换句话说, 描述一个不是 Borel 集的集合非常困难 (但也不是不可能). 我们必须不情愿地提醒读者记住：在 $\mathbb{R}$ 上存在这样的子集, 它不是 Borel 集.

定理 3 $\mathbb{R}$ 上的概率测度由它在 $(-\infty,t]$ 上的取值唯一决定. 也就是说, 如果 $\mathbb{P}$ 和 $\mathbb{Q}$ 是 $\mathbb{R}$ 上的概率测度, 且对所有的 $t\in\mathbb{R}$, 有

$$\mathbb{P}((-\infty,t])=\mathbb{Q}((-\infty,t]),$$

那么 $\mathbb{P}=\mathbb{Q}$.

证明 因为对 $s \leqslant t$, 有

$$(s,t]=(-\infty,t]\backslash(-\infty,s].$$

推得

$$\mathbb{P}((s,t])=\mathbb{Q}((s,t]).$$

因为任何开间 (s,t) 是递增的半开闭区间的并

$$(s,t)=\bigcup_{n=1}^{\infty}\left(s,t-\frac{1}{n}\right].$$

所以, 概率测度的单调连续性告诉我们

$$\mathbb{P}((s,t))=\lim_{n\to\infty}\mathbb{P}\left(\left(s,t-\frac{1}{n}\right]\right)=\lim_{n\to\infty}\mathbb{Q}\left(\left(s,t-\frac{1}{n}\right]\right)=\mathbb{Q}((s,t)).$$

因此, $\mathbb{P}$ 和 $\mathbb{Q}$ 在开区间上是相等的. 现在, 如果 $\mathbb{P}$ 和 $\mathbb{Q}$ 在开区间上是相等的, 那么可以证明它们在所有的 Borel 集上也是相等的. 然而, 我们省略这部分的证明, 因为需要一些其他的概念 (比如单调类), 这会使我们远离自己的目标.

8.3 分布函数

在有限 (或离散) 概率空间中, 概率测度基本上很容易由它们的质量函数描述

$$f(\omega)=\mathbb{P}(\{\omega\}).$$

但是, 质量函数一般不能普遍地用来描述实直线上的所有可能的概率测度, 更不用说任意的样本空间了. 因为这样, 我们需要引进概率分布函数这个概念.

定义 6 一个 (概率) 分布函数是这样的一个函数 $F:\mathbb{R}\to\mathbb{R}$, 具有下面的性质:

1) F 非减, 即

$$s<t \Rightarrow F(s)\leqslant F(t)$$

(注意有些作者用递增来表示该性质).

2) F 是右连续的, 即右极限处处存在, 且

$$\lim_{t\to a}F(t)=F(a).$$

3) F 满足

$$\lim_{t\to-\infty} F(t)=0,$$
$$\lim_{t\to\infty} F(t)=1.$$

图 1 给出了一个概率分布函数的图像. 注意该函数是非减、右连续 (但不一定连续), 在 $\pm\infty$ 处有合适的极限.

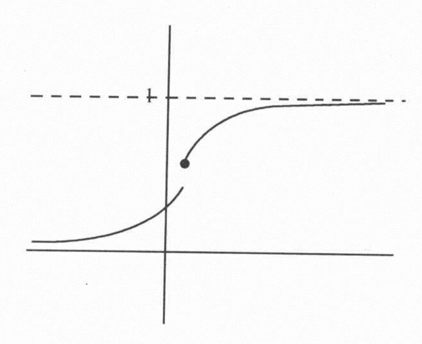


图 1 概率分布函数

概率分布函数的重要性在接下来的定理中给出. 主要地, 它暗示了 $\mathbb{R}$ 上的概率测度和概率分布函数之间存在一一对应的关系. 因此, 我们知道一个能唯一地确定另外一个, 故这两个概念在本质上是等价的. 我们省略下面定理的证明.

定理 4 1) 设 $\mathbb{P}$ 是 $\mathbb{R}$ 上的一个概率测度. 定义函数 $F_{\mathbb{P}}:\mathbb{R}\to\mathbb{R}$ 如下

$$F_{\mathbb{P}}(t)=\mathbb{P}((-\infty,t]).$$

则 $F_{\mathbb{P}}$ 是一个概率分布函数, 称为 $\mathbb{P}$ 的分布函数.

2) 设 $F:\mathbb{R}\to\mathbb{R}$ 是一个分布函数. 那么存在 $\mathbb{R}$ 上唯一的概率测度 $\mathbb{P}_F$, 使得它的分布函数是 F, 即

$$\mathbb{P}_F((-\infty,t])=F(t).$$

假设我们从一个概率测度 $\mathbb{P}$ 开始, 取它的分布函数 $F_{\mathbb{P}}$, 然后形成 $F_{\mathbb{P}}$ 的概率测度 $\mathbb{Q}$. 根据定义, 有

$$\mathbb{Q}((-\infty,t])=F_{\mathbb{P}}(t)=\mathbb{P}((-\infty,t]).$$

所以 $\mathbb{P}$ 和 $\mathbb{Q}$ 在所有的射线 $(-\infty,t]$ 上是一致的. 我们已经知道这意味着 $\mathbb{P}=\mathbb{Q}$. 因此, 从概率测度到分布函数的对应

$$\mathbb{P}\to F_{\mathbb{P}}$$

和从分布函数到概率测度的对应

$$F \to \mathbb{P}_F$$

是一一对应的, 一个为另一个的逆. 这就建立了概率测度和分布函数的概念是等价的这个概念.

例 1　$\mathbb{R}$ 上最简单的概率测度可能就是那些传达在区间 $[a,b]$ 上有“相等的可能性”或“均匀的概率”等想法的概率测度了. 例如, 考虑闭的单位区间 $[0,1]$.

我们如何表达区间 $[0,1]$ 内的每个结果有相等的可能性这个想法呢? 在有限情形, 比如说样本空间为 $\{1,\cdots,n\}$, 可简单地给每个初等事件 $\{k\}$ 分配相同的概率 $1/n$. 但是, 不像有限情形那样, 不可能给每个初等事件 $\{r\}, r \in [0,1]$ 分配一个正的实数 p, 因为这有无穷多个初等事件, 所以它们概率值的和就不会是有限的, 更不用说等于 1 了. 我们必须接受一个事实, 每个初等事件的概率为 0, 然后转向更复杂的 Borel 集.

首先, 我们观察到如果 B 是一个 Borel 集, 那么 B 中落在区间 $[0,1]$ 外的部分对概率不贡献任何力量. 换句话说,

$$\mathbb{P}(B \cap [0,1]^c) = 0.$$

至于 B 的剩余部分, 即集合 $B \cap [0,1]$, 均匀概率的想法暗示 $\mathbb{P}(B \cap [0,1])$ 应该与 $B \cap [0,1]$ 的“长度”成比例, 而不管这意味着什么.

对区间来说, 能完美地定义长度

$$\mathrm{len}([a,b]) = b - a.$$

因此, 看起来合理的概率测度的定义应该是

$$\mathbb{P}((-\infty,t]) = \mathbb{P}([0,t]) = \begin{cases} 0, & t < 0, \\ t, & 0 \leqslant t \leqslant 1, \\ 1, & t > 1. \end{cases}$$

这就是 $[0,1]$ 上的均匀分布函数. 图 2 给出了该分布函数的图像.

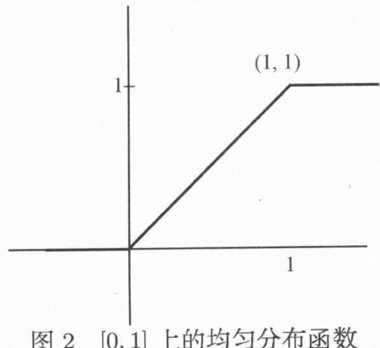


图 2　$[0,1]$ 上的均匀分布函数

例 2 正态分布是所有概率分布中最重要的一类, 它的分布函数为

$$\phi_{\mu,\sigma}(t)=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{t}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x.$$

这是一个非常复杂的函数, 但是这没有什么需要我们来处理它的. 大自然总不会用几个简单的公式而使我们的生活容易. 图 3 给出了正态分布函数的图像.

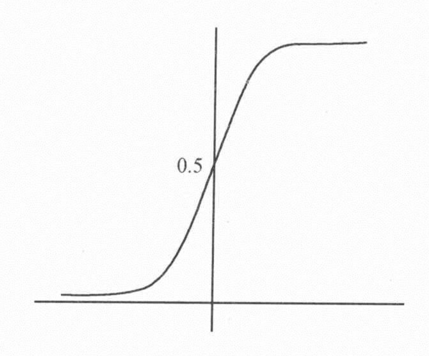


图 3 正态分布函数

参数 μ 和 σ^2 分别被称作均值 (期望值) 和方差. 标准正态分布是均值为 0 方差为 1 的正态分布, 因此它的分布函数为

$$\phi_{0,1}(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x.$$

正态分布被认为是最重要的一类分布, 其原因不仅仅是因为它在应用中经常出现. 事实上, 它为什么在应用中如此频繁地出现才是重点. 这个原因可以由概率论中最著名的定理 —— 中心极限定理从数学上给出解释. 我们会在本章的后面讨论该定理.

注意到均匀分布函数和正态分布函数都是连续的, 而不仅仅是右连续. 换句话说, 它们的曲线图没有跳跃. 分布函数中的一个跳跃表示概率测度在该点的值不为 0. 我们用一个例子来说明.

例 3 一家大众药物生产公司有一种新药正在等待食品及药物管理局的正式批准. 如果被批准, 该公司估计它该天的股价收盘价会在 $[10,15]$ 中的某个位置, 每种价格等可能地出现. 然而, 如果没有被批准, 股价就很可能是 5. 假设被批准的概率为 0.75.

我们对这种情况进行建模, 取样本空间为 $\Omega=\{5\}\cup[10,15]$, 但是用 $\mathbb{R}$ 作为样本空间可能更简单, 然后简单地将集合 Ω 外的地方的概率赋 0 值. 该概率测度的分布函数为

$$F(t)=\begin{cases}0, & t<5,\\ 0.25, & 5\leqslant t\leqslant 10,\\ 0.25+0.75\left(\dfrac{t-10}{5}\right), & 10\leqslant t<15,\\ 1, & t\geqslant 15.\end{cases}$$

图 4 给出了曲线图, 注意在 $t=5$ 处有跳跃.

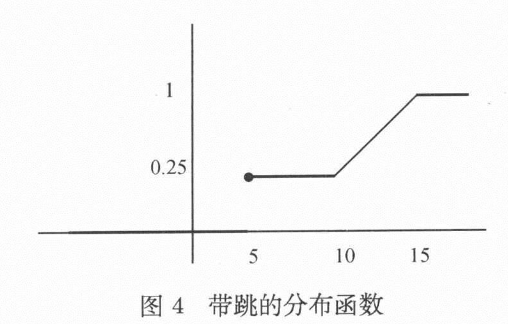


图 4　带跳的分布函数

8.4　密 度 函 数

概率测度的分布函数极其重要, 但它不一定是描述概率测度最简单的方法. 在应用中用到的许多概率测度具有一个性质, 它们的分布函数是可微的且导数“表现优美”.

说表现优美, 我们的意思是分布函数 F 的导数可积且积分又能等于 F (存在一些函数, 它的导数是可积的, 但积分不再是原函数). 用符号表示如下

$$F(t)=\int_{-\infty}^{t}F'(x)\mathrm{d}x.$$

称函数 $f(x)=F'(x)$ 为 $\mathbb{P}$ 的密度函数. 我们有一个正式的定义.

定义 7　*称概率测度 $\mathbb{P}$ 或等价地其分布函数 $F_{\mathbb{P}}$ 是绝对连续的, 如果它有一个密度函数. 密度函数是一个非负的函数 $f:\mathbb{R}\to\mathbb{R}$, 满足*

$$F_{\mathbb{P}}(t)=\int_{-\infty}^{t}f(x)\mathrm{d}x.$$

从该定义可得到

$$\mathbb{P}((a,b])=\int_{a}^{b}f(x)\mathrm{d}x.$$

换句话说, 区间 $(a,b]$ 的概率等于密度函数图像下方从 a 到 b 的面积.

注意密度函数一定非负且满足

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1.$$

即 f 整个图像与其下方的整个 x 轴所围成的面积必须等于 1. 事实上, 具有这种性质的任何非负函数 f 都是某个概率测度的密度函数.

具有密度函数的概率测度, 即绝对连续的概率测度, 是很特殊的. 例如, 它们的分布函数是连续的, 而不仅仅是右连续的. 因此, 没有一点的概率值是正的, 碰巧前面的例子就是这样.

例 4 $[0,1]$ 区间上的均匀分布函数是绝对连续的, 有密度函数

$$f(x) = \begin{cases} 0, & x \notin [0,1], \\ 1, & x \in [0,1]. \end{cases}$$

f 的图像如图 5 所示.

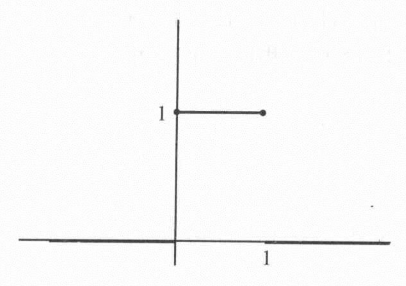


图 5 [0,1]上均匀分布的密度函数

例 5 正态分布是绝对连续的, 有密度函数

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

标准正态分布的密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}.$$

如图 6 所示. 正态密度函数的图像经常被称为钟形曲线.

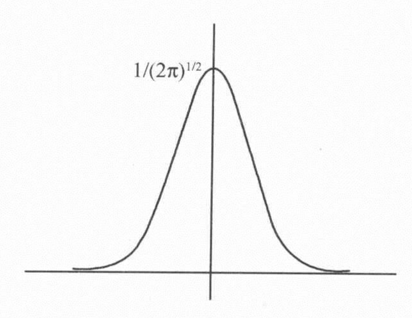


图 6 标准正态分布的密度函数

8.5　$\mathbb{R}$ 上概率测度的类型

$\mathbb{R}$ 上的概率测度可以分为如下几类: 有限、离散、绝对连续、奇异连续和混合. 我们来快速浏览每一类概率测度.

有限概率测度

称 $\mathbb{R}$ 上的概率测度 $\mathbb{P}$ 是有限的, 如果存在有限个实数 $\{r_1, \cdots, r_n\}$, 使得

$$\sum_{i=1}^{n} \mathbb{P}(r_i) = 1.$$

换句话说, 所有的概率集中在有限的点上. 有限概率测可由它的概率质量函数来描述, 概率质量函数处处为零, 除了一些取正概率的点

$$f(x) = \begin{cases} \mathbb{P}(r_i), & x = r_i, \\ 0, & \text{其他情况}. \end{cases}$$

也被称为 $\mathbb{P}$ 的密度函数. 有限概率测度的分布函数有有限个跳跃点, 在其他地方为常数. 如图 7 所示.

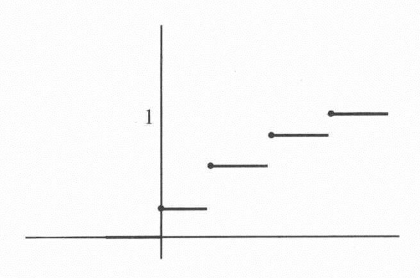


图 7　有限概率测度的分布函数

离散概率测度

称 $\mathbb{R}$ 上的概率测度 $\mathbb{P}$ 是离散的, 如果存在可数个实数 $\{r_1, r_2, \cdots\}$, 使得

$$\sum_{i=1}^{\infty} \mathbb{P}(r_i) = 1.$$

换句话说, 所有的概率集中在一列可数的点上 (说可数, 意思是有限或者可数无限. 因此, 有限概率测度也是离散概率测度). 如同有限概率测度那样, 一般的离散概率

测度可由概率质量函数来描述, 虽然在这种情形下, "密度函数" 这个术语被更广泛地使用.

$$f(x) = \begin{cases} \mathbb{P}(r_i), & x = r_i, \\ 0, & \text{其他情况}. \end{cases}$$

实际上, 离散概率测度的分布函数很复杂. 当函数确实只有可数个跳跃点时, 这些点可能发生在如有理数集之类的集合上, 而所有的有理数 "均匀" 地散布在整个实直线上.

绝对连续的概率测度

如我们已经看到的, 称 $\mathbb{R}$ 上的概率测度 $\mathbb{P}$ 是绝对连续的, 如果它有一个密度函数, 即一个非负的函数 $f:\mathbb{R}\to\mathbb{R}$, 满足

$$F_{\mathbb{P}}(t) = \int_{-\infty}^{t} f(x)\mathrm{d}x.$$

奇异连续的概率测度

奇异连续的概率测度在本质上是很反常的. 奇异连续的概率测度是这样的一个概率测度, 它的分布函数是可微的 (因此是连续的), 但是导数在整个实直线上几乎处处为 0(除了一个概率为零的集合). 幸运的是, 在本书中我们不需要处理这类概率测度.

混合的概率测度

实际上 $\mathbb{R}$ 上的任何概率测度 $\mathbb{P}$ 可以分解为 (唯一的方式) 一个离散概率测度, 一个绝对连续概率测度和一个奇异连续的概率测度的线性组合, 用符号表示为

$$\mathbb{P} = \alpha_d\mathbb{P}_d + \alpha_a\mathbb{P}_a + \alpha_s\mathbb{P}_s,$$

其中系数 α_d, α_a 和 α_s 是非负的, 且满足 $\alpha_d + \alpha_a + \alpha_s = 1$. 因此, 所有的概率测度或者是离散的、绝对连续的、奇异连续的, 或者是这些类型的凸组合.

8.6 随机变量

正如非有限情形中的事件是比较复杂的那样, 随机变量这个概念也是. 特别地, 不是所有的函数都是随机变量.

定义 8 设 (Ω, Σ) 是一个可测空间. 称函数 $X : \Omega \to \mathbb{R}$ 是 Σ 可测的, 如果每个开区间的原象都在 Σ 中, 用符号表示就是

$$X^{-1}((a, b)) \in \Sigma.$$

称 (Ω, Σ) 上的可测函数为随机变量.

该定义说明一个随机变量 X 具有性质: 集合 $X^{-1}((a, b))$ 必须是"可测的".

下面是有关随机变量的一些事实, 其证明就省略了.

定理 5 1) 随机变量的和或者乘积也是随机变量, 从而一个随机变量乘以任何常数还是一个随机变量.

2) 随机变量的复合也是随机变量.

3) 连续和点连续的函数是随机变量.

随机变量的分布函数

如果 $(\Omega, \Sigma, \mathbb{P})$ 是一个任意的样本空间, X 是 (Ω, Σ) 上的一个随机变量, 那么 X 按如下方式定义了 $\mathbb{R}$ 上的一个分布函数 F_X 和一个相应的概率测度 $\mathbb{P}_X$,

$$F_X(t) = \mathbb{P}_X((-\infty, t]) = \mathbb{P}(X \leqslant t).$$

可证明函数 $F(t) = \mathbb{P}(X \leqslant t)$ 是一个分布函数. 如果 $\mathbb{P}_X$ 是有限的、离散的或绝对连续的, 那么我们相应地说随机变量是有限的、离散的或绝对连续的. 为了简单, 经常称绝对连续的随机变量为连续随机变量.

由随机变量生成的 σ 代数

如果 $X : (\Omega, \Sigma) \to \mathbb{R}$ 是一个随机变量, 那么任何 Borel 集的原象都在 Σ 里. 但是, 不要求 Σ 的所有元素都是 X 的原象. X 在 Σ 中的原象形成了另外一个 σ 代数, 它是 Σ 的一个子 σ 代数.

定义 9 由随机变量 $X : (\Omega, \Sigma) \to \mathbb{R}$ 生成的 σ 代数是这样的一个 σ 代数 $\sigma(X)$, 其元素为 $\mathbb{R}$ 上的 Borel 集在 X 下的原象, 即

$$\sigma(X) = \{\{X \in B\} | \ B \in \mathcal{B}\}.$$

σ 代数 $\sigma(X)$ 有一个独特的性质, 即它是 Ω 上使得 X 可测的最小 σ 代数. 非严格地说, 它是刚好使 X 可测所需要的 σ 代数. 下面的定理与可测性的定义没什么区别.

定理 6 设 $X : \Omega \to \mathbb{R}$ 是一个函数, 令 Σ 是 Ω 上的 σ 代数. 那么 X 是 Ω 可测的当且仅当 Σ 包含 $\sigma(X)$.

随机变量的独立性

下面给出了任意随机变量相互独立的定义.

定义 10 称 $\mathbb{R}$ 上的两个随机变量 X 和 Y 相互独立, 如果对所有的 $s,t \in \mathbb{R}$, 有

$$\mathbb{P}(X \leqslant t, Y \leqslant s) = \mathbb{P}(X \leqslant t)\mathbb{P}(Y \leqslant s).$$

更一般地, 称一列随机变量 $X_1, \cdots, X_n$ 相互独立, 如果

$$\mathbb{P}(X_1 \leqslant t_1, \cdots, X_n \leqslant t_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \leqslant t_i).$$

该定义正式地表达了这样一种感觉, 如果随机变量相互独立, 那么一个随机变量的取值不影响另外那些随机变量的取值.

随机变量的期望和方差

回忆一下有限概率空间 $(\Omega, \mathbb{P})$ 上的随机变量 X, 这里 $\Omega = \{\omega_1, \cdots, \omega_n\}$, 它的期望 (均值) 定义如下

$$\mathcal{E}_{\mathbb{P}}(X) = \sum_{i=1}^{n} X(\omega_i)\mathbb{P}(\omega_i).$$

如果 $g : \mathbb{R} \to \mathbb{R}$ 是一个函数, 那么随机变量 $g(X)$ 的期望为

$$\mathcal{E}_{\mathbb{P}}(g(X)) = \sum_{i=1}^{n} g(X(\omega_i))\mathbb{P}(\omega_i).$$

同样, 方差的定义为

$$\mathrm{var}(X) = \mathcal{E}((x-\mu)^2).$$

我们现在将这些概念扩展到绝对连续随机变量上去.

定义 11 设 X 是一个绝对连续的随机变量, 有密度函数 f. X 的期望或均值就是广义积分

$$\mathcal{E}(X) = \int_{-\infty}^{\infty} x f(x)\mathrm{d}x.$$

该积分存在, 如果有

$$\int_{-\infty}^{\infty} |x| f(x)\mathrm{d}x < \infty.$$

X 的方差为

$$\mathrm{var}(X) = \mathcal{E}((X-\mu)^2).$$

标准差为方差的正平方根

$$SD(X) = \sqrt{\text{var}(X)}.$$

同样, 如果 $g : \mathbb{R} \to \mathbb{R}$ 是一个可测函数, 那么随机变量 $g(X)$ 的期望为

$$\mathcal{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x,$$

考虑到如果有

$$\int_{-\infty}^{\infty} |g(x)|f(x)\mathrm{d}x < \infty.$$

下面为期望和方差的一些基本性质.

定理 7　1) 期望算子是线性的, 即

$$\mathcal{E}(aX + bY) = a\mathcal{E}(X) + b\mathcal{E}(Y).$$

2) 如果 $X_1, \cdots, X_n$ 是 $\mathbb{R}$ 上相互独立的随机变量, 那么

$$\mathcal{E}(X_1 \cdots X_n) = \prod_{i=1}^{n} \mathcal{E}(X_i).$$

3) $\text{var}(X) = \mathcal{E}(X^2) - \mu^2 = \mathcal{E}(X^2) - \mathcal{E}(X)^2$.

4) 对任意实数 a, 有

$$\text{var}(aX) = a^2\text{var}(X)$$

和

$$\text{var}(X - a) = \text{var}(X).$$

5) 如果 $X_1, \cdots, X_n$ 是 $\mathbb{R}$ 上相互独立的随机变量, 那么

$$\text{var}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \text{var}(X_i).$$

8.7 正态分布

我们换一个角度来看看正态分布, 其密度函数为

$$N_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

参数 μ 和 σ^2 分别是均值和方差. 为了计算均值, 我们只需要一点一年级的微积分. 均值的定义是

$$\mathcal{E} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x.$$

将 x 写成 $(x-\mu)+\mu$, 拆开积分得到

$$\mathcal{E}=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}(x-\mu)\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x+\frac{\mu}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x.$$

第二部分积分等于 μ 乘以 $N_{\mu,\sigma}$ 的积分, 因为 $N_{\mu,\sigma}$ 的积分等于 1, 所以我们得到该部分为 μ. 对于第一部分积分, 做替换 $y=x-\mu$ 得到

$$\int_{-\infty}^{\infty}y\mathrm{e}^{-\frac{y^2}{2\sigma^2}}\mathrm{d}y.$$

由于被积函数是一个奇函数, 因而从 $-\infty$ 到 ∞ 的积分一定为 0 (我们将细节留作练习). 因此, $\mathcal{E}=0+\mu=\mu$.

计算正态分布的方差需要用到一个漂亮但非显然的公式

$$\int_{-\infty}^{\infty}y^2\mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y=\sqrt{2\pi}.$$

从这里开始, 剩下的就是直接计算, 特别地用到公式

$$\mathrm{var}(X)=\mathcal{E}(X^2)-\mathcal{E}(X)^2.$$

得到的结果是 $\mathrm{var}(N_{\mu,\sigma})=\sigma^2$. 将细节留作练习.

假设 $\mathcal{N}_{\mu,\sigma}$ 是均值为 μ 方差为 σ^2 的正态随机变量. 考虑随机变量

$$Z=\frac{\mathcal{N}_{\mu,\sigma}-\mu}{\sigma}.$$

根据期望和方差的性质有

$$\mathcal{E}(Z)=\frac{1}{\sigma}\mathcal{E}(\mathcal{N}_{\mu,\sigma}-\mu)=\frac{1}{\sigma}(\mathcal{E}(\mathcal{N}_{\mu})-\mu)=0$$

和

$$\mathrm{var}(Z)=\frac{1}{\sigma^2}\mathrm{var}(\mathcal{N}_{\mu,\sigma}-\mu)=\frac{1}{\sigma^2}\mathrm{var}(\mathcal{N}_{\mu,\sigma})=1.$$

接下来计算 Z 的分布, 有

$$\begin{aligned}\mathbb{P}(Z\leqslant t)&=\mathbb{P}\left(\frac{\mathcal{N}_{\mu,\sigma}-\mu}{\sigma}\leqslant t\right)\\&=\mathbb{P}(\mathcal{N}_{\mu,\sigma}\leqslant\sigma t+\mu)\\&=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\sigma t+\mu}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x.\end{aligned}$$

作替换 $y=(x-\mu)/\sigma$, 得到

$$\mathbb{P}(Z\leqslant t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}\mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y.$$

所以 $Z=\mathcal{N}_{0,1}$ 是标准正态随机变量. 称从 $\mathcal{N}_{\mu,\sigma}$ 到 Z 的过程为标准化过程.

定理 8　如果 $\mathcal{N}_{\mu,\sigma}$ 是均值为 μ 方差为 σ^2 的正态随机变量, 那么

$$\mathcal{N}_{0,1} = \frac{\mathcal{N}_{\mu,\sigma} - \mu}{\sigma}$$

是一个标准正态随机变量. 同样地, 如果 $\mathcal{N}_{0,1}$ 是一个标准正态随机变量, 那么

$$\mathcal{N}_{\mu,\sigma} = \sigma\mathcal{N}_{0,1} + \mu$$

是均值为 μ 方差为 σ^2 的正态随机变量.

我们将用到一个与正态分布有关的分布, 该分布就是对数正态分布. 如果一个随机变量 X 具有性质: $\log X$ 是正态分布. 那么称 X 是对数正态分布 (注意, X 是对数正态的如果它的对数是正态的, 而不是说它是一个正态随机变量的对数. 换句话说, 对数正态意味着"对数是正态的"而不是"正态的对数").

下面的证明留作练习.

定理 9　如果 X 服从对数正态分布, 比如说 $Y = \log X$ 是均值为 a 方差为 b^2 的正态分布, 那么

$$\begin{aligned}\mathcal{E}(X) &= \mathcal{E}(\mathrm{e}^Y) = \mathrm{e}^{a+\frac{1}{2}b^2},\\ \mathrm{var}(X) &= \mathrm{var}(\mathrm{e}^Y) = \mathrm{e}^{2a+b^2}(\mathrm{e}^{b^2} - 1).\end{aligned}$$

8.8　依分布收敛

下面是点收敛的定义.

定义 12　设 (f_n) 是 $\mathbb{R}$ 到 $\mathbb{R}$ 的函数序列, f 是另外一个函数. 如果对每个实数 x, 有实数序列 $(f_n(x))$ 收敛到 $f(x)$, 那么称 (f_n) 点收敛到 f.

如果熟悉普通实数序列的收敛, 那么实际上对函数的点收敛也会感到熟悉. 这里没有什么新的东西.

现在考虑一列随机变量序列 (X_n). 当然, 这些随机变量也是函数, 虽然是特殊的函数. 设 X 是另外一个随机变量. 结果表明, 存在几种有用的方式可用来定义 (X_n) 到 X 的收敛 (当中只有一种是点收敛). 但是, 我们对其中的一种收敛感兴趣. 下面就是它的定义.

定义 13　设 (X_n) 是一列随机变量, 这里允许每个随机变量可以定义在不同的概率空间 $(\Omega_n, \mathbb{P}_n)$ 上. 设 X 是概率空间 $(\Omega, \mathbb{P})$ 上的随机变量. 那么称 (X_n) 依分布收敛到 X, 写成

$$X_n \xrightarrow{\text{dist}} X,$$

如果在 F_X 的所有连续点上, 分布函数 (F_{X_n}) 依点收敛到分布函数 F_X. 因此, 如果 F_X 在点 s 处连续, 那么一定有

$$\lim_{n\to\infty} F_{X_n}(s) = F_X(s).$$

也就是

$$\lim_{n\to\infty} \mathbb{P}_n(X_n \leqslant s) = \mathbb{P}(X \leqslant s).$$

依分布收敛也叫做弱收敛.

关于弱收敛, 我们需要下面的结果.

定理 10 设 (X_n) 是一列随机变量, 其中 X_n 定义在 $(\Omega_n, \mathbb{P}_n)$ 上. 又设 X 是概率空间 $(\Omega, \mathbb{P})$ 上的随机变量.

1) 对所有的有界连续函数 $g : \mathbb{R} \to \mathbb{R}$, 有

$$X_n \xrightarrow{\text{dist}} X \text{ 当且仅当 } \mathcal{E}_{\mathbb{P}_n}(g(X_n)) \to \mathcal{E}_{\mathbb{P}}(g(X)).$$

特别地,

$$\mathcal{E}_{\mathbb{P}_n}(X_n) \to \mathcal{E}_{\mathbb{P}}(X).$$

2) 对所有的连续函数 $f : \mathbb{R} \to \mathbb{R}$, 有

$$f(X_n) \xrightarrow{\text{dist}} f(X).$$

证明 我们省略 1) 的证明. 对于 2), 设 f 连续. 那么对任何有界连续函数 g, 复合函数 $g \circ f$ 也是有界连续的. 因此, 由 1) 得

$$\mathcal{E}(g(f(X_n))) = \mathcal{E}((g \circ f)(X_n)) \to \mathcal{E}((g \circ f)(X)) = \mathcal{E}(g(f(X))).$$

由 1) 知

$$f(X_n) \xrightarrow{\text{dist}} f(X).$$

得证.

定理 11 设 (X_n) 是一列随机变量, 且有

$$X_n \xrightarrow{\text{dist}} X,$$

其中 X 是一个分布函数连续的随机变量. 如果 (a_n) 和 (b_n) 是实数列, 且有

$$a_n \to a, \quad b_n \to b.$$

那么

$$a_n X_n + b_n \xrightarrow{\text{dist}} aX + b.$$

特别地, 如果 $a \neq 0$ 且 $X_n \xrightarrow{\text{dist}} \mathcal{N}_{0,1}$, 其中 $\mathcal{N}_{0,1}$ 是标准正态随机变量, 那么

$$a_n X_n + b_n \xrightarrow{\text{dist}} \mathcal{N}_{a,b},$$

其中 $\mathcal{N}_{a,b}$ 是均值为 a 方差为 b^2 的正态随机变量.

证明　下面的证明需要一致收敛的概念. 如果对这些概念不熟悉, 读者可跳过该证明. 令 F_{X_n} 和 F_X 分别表示 X_n 和 X 的分布函数.

第一步就是证明对任何 $s \in \mathbb{R}$, 存在一个区间 $(s-\lambda, s+\lambda)$, 使得 F_{X_n} 在该区间上一致收敛到 F_X. 为了证明此, 我们利用分布函数是非减的这个事实. 所以给定 $\varepsilon > 0$, 有

$$\begin{aligned} F_{X_n}(t) - F_X(t) &\leqslant F_{X_n}(s+\alpha) - F_X(s-\alpha) \\ &= [F_{X_n}(s+\alpha) - F_X(s+\alpha)] + [F_X(s+\alpha) - F_X(s-\alpha)]. \end{aligned}$$

因为 F_{X_n} 依点收敛到 F_X, 所以可以选择一个 α_1, 使得

$$F_{X_n}(s+\alpha_1) - F_X(s+\alpha_1) < \varepsilon/2.$$

而且, 因为 F_X 在 s 处连续, 从而可以选择一个 α_2, 使得

$$F_X(s+\alpha) - F_X(s-\alpha) < \varepsilon/2.$$

因此, 取 $\alpha = \min\{\alpha_1, \alpha_2\}$, 对所有 $t \in (s-\alpha, s+\alpha)$ 有

$$F_{X_n}(t) - F_X(t) < \varepsilon.$$

另一方面, 我们也有

$$\begin{aligned} F_{X_n}(t) - F_X(t) &\geqslant F_{X_n}(s-\beta) - F_X(s+\beta) \\ &= [F_{X_n}(s-\beta) - F_X(s-\beta)] + [F_X(s-\beta) - F_X(s+\beta)]. \end{aligned}$$

很明显可以选取一个 β_1, 使得

$$F_{X_n}(s-\beta_1) - F_X(s-\beta_1) > -\varepsilon/2$$

和 β_2, 使得

$$F_X(s-\beta_2) - F_X(s+\beta_2) > -\varepsilon/2.$$

因此, 取 $\beta = \min\{\beta_1, \beta_2\}$, 对所有 $t \in (s-\beta, s+\beta)$, 有

$$-\varepsilon < F_{X_n}(t) - F_X(t).$$

最后, 令 λ 为 α 和 β 的较小者, 对所有的 $t \in (s-\lambda, s+\lambda)$, 有

$$-\varepsilon < F_{X_n}(t) - F_X(t) < \varepsilon.$$

这就证明了 F_{X_n} 在 $(s-\lambda, s+\lambda)$ 上一致收敛到 F_X.

现在我们利用该结果. 令 $t \in \mathbb{R}$, 选取一个 λ, 使得 F_{X_n} 在下面的区间上一致收敛到 F_X,

$$I = \left(\frac{t-b}{a} - \lambda, \frac{t-b}{a} + \lambda\right).$$

对任意的 $\varepsilon > 0$, 存在 N_1, 使得对所有的 $s \in I$, 有

$$n > N_1 \Rightarrow |F_{X_n}(s) - F_X(s)| < \frac{\varepsilon}{2}.$$

同样, 存在 $N_2 > 0$, 使得

$$n > N_2 \Rightarrow \frac{t-b_n}{a_n} \in I.$$

从而

$$n > \max\{N_1, N_2\} \Rightarrow \left|F_{X_n}\left(\frac{t-b_n}{a_n}\right) - F_X\left(\frac{t-b_n}{a_n}\right)\right| < \frac{\varepsilon}{2}.$$

同样地, F_X 的连续性意味着存在 $N_3 > 0$, 使得

$$n > N_3 \Rightarrow \left|F_X\left(\frac{t-b_n}{a_n}\right) - F_X\left(\frac{t-b}{a}\right)\right| < \frac{\varepsilon}{2}.$$

因此

$$n > \max\{N_1, N_2, N_3\} \Rightarrow \left|F_{X_n}\left(\frac{t-b_n}{a_n}\right) - F_X\left(\frac{t-b}{a}\right)\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

但是

$$F_{X_n}\left(\frac{t-b_n}{a_n}\right) = \mathbb{P}\left(X_n \leqslant \frac{t-b_n}{a_n}\right) = \mathbb{P}(a_n X_n + b_n \leqslant t),$$

并且

$$F_X\left(\frac{t-b}{a}\right) = \mathbb{P}\left(X \leqslant \frac{t-b}{a}\right) = \mathbb{P}(aX + b \leqslant t).$$

所以我们证明了

$$\mathbb{P}(a_n X_n + b_n \leqslant t) \to \mathbb{P}(aX + b \leqslant t),$$

即

$$a_n X_n + b_n \xrightarrow{\text{dist}} aX + b.$$

由 $a\mathcal{N}_{0,1} + b = \mathcal{N}_{a,b}$ 可得证第二部分.

8.9　中心极限定理

中心极限定理是概率论中最著名的定理, 这是有很好的理由的. 实际上, 中心极限定理有好几个版本. 我们首先陈述最常见的版本, 然后讨论另一个不同的版本, 在下章将用到它.

从直观上看, 如果 X 是一个随机变量, 那么它的分布函数描述了它的概率“行为”或“特征”. 更准确地说, 如果 X 和 Y 是分布函数相同的随机变量, 那么对所有的实数 a 和 b, 有

$$\mathbb{P}(a \leqslant X \leqslant b) = \mathbb{P}(a \leqslant Y \leqslant b).$$

(注意, 不需要函数 X 和 Y 是相同的. 事实上, 甚至不需要它们是定义在相同的样本空间上). 当两个或更多的随机变量有相同的分布函数时, 我们称它们是同分布的.

中心极限定理指出：如果 S_n 是 n 个随机变量的和, 这些随机变量满足：

1) 相互独立.

2) 同分布.

如果将 S_n 标准化, 那么标准化后的随机变量 S_n^* 有非常特殊的特征. 特别地, S_n^* 的分布函数逼近标准正态分布函数, 而不管原随机变量是什么类型. 而且, 逼近随着 n 的变大而变得越来越好.

因此, 求和以及标准化的过程“洗掉”了单个随机变量的原有特征, 用标准正态随机变量的特征来代替了以前的特征.

下面是中心极限定理的一个正式陈述.

定理 12(中心极限定理)　设 $X_1, X_2, \cdots$ 是一列独立同分布的随机变量, 有有限均值 μ 和有限方差 $\sigma^2 > 0$. 令

$$S_n = \sum_{i=1}^{n} X_i$$

为前 n 个随机变量的和. 因此, $\mathcal{E}(S_n) = n\mu, \mathrm{var}(S_n) = n\sigma^2$. 考虑标准化了的随机变量

$$S_n^* = \frac{S_n - \mathcal{E}(S_n)}{\sqrt{\mathrm{var}(S_n)}} = \frac{S_n - n\mu}{\sqrt{n}\sigma}.$$

则标准化序列 S_n^* 依分布收敛到一个标准正态随机变量 $\mathcal{N}_{0,1}$, 即

$$\lim_{n\to\infty} F_{S_n^*}(t) = \phi_{0,1}(t).$$

换一种方式说

$$\mathbb{P}(S_{n^*} < t) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \mathrm{e}^{-\frac{x^2}{2}} \mathrm{d}x,$$

其中当 $n \to \infty$ 时, 逼近误差趋于零.

可能如你所想的那样, 中心极限定理的证明有点棘手, 我们在本书中就不探究了. 但是, 建议读者稍微停下来想一下该定理中有点令人吃惊的性质. 当然它也说明了正态分布的极其重要性.

如先前提到的那样, 我们需要中心极限定理的另一不同版本, 在布莱克 – 舒尔斯公式中要用到它. 一方面, 我们只需考虑均值为 0 方差为 1 的伯努利随机变量, 它是有用的随机变量中最简单的一类. 另一方面, 我们把事情弄得更复杂, 因为这些伯努利随机变量不是同分布的!

特别地, 我们想要考虑的不仅是简单的一列随机变量, 而且是随机变量的一个三角排列

$$\begin{array}{llll} B_{1,1} & & & \\ B_{2,1} & B_{2,2} & & \\ B_{3,1} & B_{3,2} & B_{3,3} & \\ \vdots & \vdots & \vdots & \ddots \end{array}$$

每一行的随机变量是独立同分布的伯努利随机变量, 均值为 0, 方差为 1. 特别地, $B_{n,i}$ 是一个伯努利随机变量, 有

$$\mathbb{P}\left(B_{n,i} = \frac{q_n}{\sqrt{p_n q_n}}\right) = p_n,$$

$$\mathbb{P}\left(B_{n,i} = \frac{-p_n}{\sqrt{p_n q_n}}\right) = q_n,$$

其中 $q_n = 1 - p_n$. 但是, 不同行的随机变量既不需要独立也不必同分布. 事实上, 它们甚至不必定义在相同的概率空间上. 这一点在以后对我们非常重要.

也必须假定概率 p_n 是“表现完美的”, 意思是它们不会接近 0 或 1. 事实上, 我们假定存在一个 p 满足 $0 < p < 1$, 使得

$$p_n \to p.$$

从而也有

$$q_n \to q = 1 - p \in (0, 1).$$

现在, 中心极限定理的某个版本刚好表达了这种情况 (即使当随机变量不是伯努利分布时).

我们从每个随机变量的标准化开始, 使得它的均值为 0, 每行的方差和为 1. 因此

$$\mathcal{E}(B_{n,i})=0,$$

$$\mathrm{var}(B_{n,i})=1.$$

标准化的随机变量为

$$B^*_{n,i}=\frac{1}{\sqrt{n}}B_{n,i}.$$

新排列为

$$\begin{array}{llll}
B_{1,1} & & & \\
\frac{1}{\sqrt{2}}B_{2,1} & \frac{1}{\sqrt{2}}B_{2,2} & & \\
\frac{1}{\sqrt{3}}B_{3,1} & \frac{1}{\sqrt{3}}B_{3,2} & \frac{1}{\sqrt{3}}B_{3,3} & \\
\vdots & \vdots & \vdots & \ddots
\end{array}$$

现在, 诠释了这种情况的中心极限定理指出, 在一定的条件下, 每行之和的分布

$$\begin{aligned}
S_1&=B_{1,1},\\
S_2&=\frac{1}{\sqrt{2}}(B_{2,1}+B_{2,2}),\\
S_3&=\frac{1}{\sqrt{3}}(B_{3,1}+B_{3,2}+B_{3,3}).
\end{aligned}$$

依点收敛到标准正态分布 $\phi_{0,1}$.

所谓的一定条件有点杂乱. 从直观上来说, S_n 中的每项相对于整个和来说是"可忽略的". 在我们感兴趣的伯努利随机变量情形中, 在 S_n 中出现的标准化伯努利随机变量 $B^*_{n,i}$ 的可能值为

$$\frac{q_n}{\sqrt{np_nq_n}} \quad 和 \quad \frac{-p_n}{\sqrt{np_nq_n}}.$$

现在当 $n\to\infty$ 时, 有

$$\frac{q_n}{\sqrt{p_nq_n}}\to\frac{q}{\sqrt{pq}},$$

$$\frac{-p_n}{\sqrt{p_n q_n}} \to \frac{-p}{\sqrt{pq}}.$$

因为 p 和 q 都是正的, 所以这些极限有限. 因此, $B^*_{n,i}$ 的可能值满足

$$\frac{q_n}{\sqrt{np_n q_n}} \to 0,$$

$$\frac{-p_n}{\sqrt{np_n q_n}} \to 0.$$

这是应用中心极限定理的一个充分条件. 我们最终得到了所需要的定理.

定理 13　考虑随机变量的一个三角排列

$$\begin{matrix} B_{1,1} & & & \\ B_{2,1} & B_{2,2} & & \\ B_{3,1} & B_{3,2} & B_{3,3} & \\ \vdots & \vdots & \vdots & \ddots \end{matrix}$$

其中对行 n 和 $1 \leqslant i \leqslant n$ 来说, $B_{n,i}$ 是独立同分布的伯努利随机变量, 且

$$\mathbb{P}\left(B_{n,i} = \frac{q_n}{\sqrt{p_n q_n}}\right) = p_n,$$
$$\mathbb{P}\left(B_{n,i} = \frac{-p_n}{\sqrt{p_n q_n}}\right) = q_n.$$

但是, 不同行的随机变量不需要独立或同分布, 甚至不必定义在相同的概率空间上. 同时假设 $p_n \to p \in (0,1)$. 那么"标准化"的随机变量

$$S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} B_{n,i}$$

依分布收敛到一个标准正态随机变量. 更具体地说, 如果 Z 是任意概率空间上的一个标准正态随机变量, 那么 S_n 依分布收敛到 Z.

正如在前面已经提到的那样, 我们在下一章会利用该定理来帮助推导布莱克–舒尔斯期权定价公式.

练　习　8

1. 设 $f(t)$ 是分段线性的概率密度函数, 具有性质：当 $t \leqslant 0$ 或者 $t \geqslant 2$ 时, $f(t) = 0$; $f(1) = a$. 画个密度函数的草图并求出 a. 画出相应分布函数的草图.

2. 设 X 的分布函数 F 如下

$$F(t)=\begin{cases}0, & t<0,\\ \frac{1}{2}t, & 0\leqslant t\leqslant 2,\\ 1, & t>2.\end{cases}$$

令 $Y=X^2$. 试找出

a) $\mathbb{P}(0\leqslant X\leqslant 1)$.

b) $\mathbb{P}(1\leqslant X\leqslant 3)$.

c) $\mathbb{P}(Y\leqslant X)$.

d) $\mathbb{P}\left(X+Y\leqslant \frac{3}{4}\right)$.

e) 随机变量 $\sqrt{X}$ 的分布函数.

3. 设 $\Omega=\{\omega_1,\omega_2,\omega_3\}$, 令 $\mathbb{P}$ 是 Ω 上的均匀概率测度, 即 $\mathbb{P}(\omega_i)=1/3, i=1,2,3$. 考虑如下随机变量

$$\begin{aligned}&X(\omega_1)=1, \quad &&X(\omega_2)=2, \quad &&X(\omega_3)=3,\\ &Y(\omega_1)=2, &&Y(\omega_2)=3, &&Y(\omega_3)=1,\\ &Z(\omega_1)=3, &&Z(\omega_2)=1, &&Z(\omega_3)=2.\end{aligned}$$

这些随机变量是一样的吗? 它们的分布函数是什么?

4. 证明 σ 代数可数交运算是封闭的.

5. 证明所有的半直线都是 Borel 集.

6. 证明所有的闭区间都是 Borel 集.

7. 证明：对任何事件 A 和 B, 有

$$\mathbb{P}(A\cup B)+\mathbb{P}(A\cap B)=\mathbb{P}(A)+\mathbb{P}(B).$$

8. 证明概率测度是次可加的, 即对任何事件 A 和 B, 有

$$\mathbb{P}(A\cup B)\leqslant \mathbb{P}(A)+\mathbb{P}(B).$$

9. 找出区间 $[a,b]$ 上的均匀分布函数并作草图.

10. 证明最一般的伯努利随机变量 B, 其均值为 0, 方差为 1, 由如下给出

$$\mathbb{P}\left(B=\frac{q}{\sqrt{pq}}\right)=p,$$

$$\mathbb{P}\left(B=\frac{-p}{\sqrt{pq}}\right)=q,$$

其中 $q=1-p$.

11. 完成正态分布均值为 μ 的证明中的细节部分.

12. 利用文中的积分公式计算正态分布的方差.

13. 证明：如果 $A_1 \subseteq A_2 \subseteq \cdots$ 是一列递增的事件序列, 每个事件都被下一个事件包含, 那么 $\lim\limits_{i\to\infty} \mathbb{P}(A_i)$ 存在, 且

$$\lim_{i\to\infty} \mathbb{P}(A_i) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right).$$

14. 设 X 是 $\mathbb{R}$ 上的随机变量. 证明函数

$$f(t) = \mathbb{P}(X \leqslant t)$$

是一个概率分布函数. 提示：充分利用单调连续性.

15. 验证概率测度 $\mathbb{P}$ 和它的分布函数 F 之间有

a) $\mathbb{P}((a,b]) = F(b) - F(a)$.

b) $\mathbb{P}((a,b)) = F(b-) - F(a)$.

c) $\mathbb{P}([a,b)) = F(b-) - F(a-)$.

d) $\mathbb{P}([a,b]) = F(b) - F(a-)$.

这里负号表示从下取极限.

16. 如果 $X_n \xrightarrow{\text{dist}} X$, 那么对任何实数 $a \neq 0$ 和 b, 有

$$aX_n + b \xrightarrow{\text{dist}} aX + b.$$

第 9 章　布莱克 – 舒尔斯期权定价公式

我们已经学过的模型都是离散时间模型, 因为在这些模型中, 从时间上看改变仅发生在离散点上. 另一方面, 在连续模型中, 改变可能发生在 (至少理论上是) 模型周期内的任何时刻.

最著名的连续时间衍生品定价模型就是布莱克 – 舒尔斯期权定价模型 (Black-Scholes). 基于下面的五个量, 该模型给出了欧式期权的价格：

- 标的股票的初始价格, 该价格是已知的.
- 期权的执行价, 该价格也是已知的.
- 到期日, 也是已知的.
- 期权有效期内的无风险利率, 假定该利率为常数并只能估计得出.
- 股票价格的波动率, 为一个常数. 它度量了股价的波动幅度, 因此也是该股票的风险的一个度量. 这个量也只能通过估计得出.

在这一章中, 我们的目标就是刻画连续时间模型和推导布莱克 – 舒尔斯期权定价公式. 我们将从 CRR 模型的极限情况来推导连续时间模型.

9.1　股票价格和布朗运动

1827 年, 刚好是纽约股票交易所成立 35 年后, 一位名叫布朗 (Brown) 的英国植物学家研究了液体中的小花粉颗粒的运动情况. 他写道, 在显微镜下观察到花粉颗粒的运动是连续无规则的.

爱因斯坦在 1905 年第一次对该现象给出了科学的解释. 他指出这种无规则的运动 (现在被称为布朗运动) 是因为花粉颗粒受到液体分子的连续强烈撞击而产生的结果. 布朗运动在数学上的正式描述和性质最早是由伟大的数学家罗伯特 · 维纳 (Robert Wiener) 在 1918 年给出的.

对我们来说特别有趣的是, 1900 年法国数学家巴舍利耶 (Bachelier) 在他的博士论文中用布朗运动来刻画股票价格的运动.

布朗运动 (Brown Motion)

现在来近距离看看布朗运动. 我们已经定义了有限随机过程就是样本空间 Ω 上的一列随机变量 $X_1, \cdots, X_N$. 在区间 $I \subseteq \mathbb{R}$ 上的连续随机过程就是 Ω 上的随机

变量的群集 $\{X_t|\ t \in I\}$. 一般取 $[0,\infty)$ 作为 I. 变量 t 通常代表时间, 因此随机过程在时间 t 的值就是随机变量 X_t 的取值.

定义 1 称一个连续随机过程 $\{W_t|\ t \geqslant 0\}$ 是一个波动率为 σ 的布朗运动或者维纳过程, 如果

1) $W_0 = 0$.

2) W_t 是均值为 0 方差为 $\sigma^2 t$ 的正态分布.

3) 过程 $\{W_t\}$ 具有平稳增量性质, 也就是说对任意 $s < t$, 增量 $W_t - W_s$ 只依赖 $t-s$. 因此, $W_t - W_s$(与 $W_{t-s} - W_0 = W_{t-s}$ 是同分布的) 是均值为 0 方差为 $\sigma^2(t-s)$ 的正态分布.

4) 过程 $\{W_t\}$ 具有独立增量性质, 也就是说, 对任意时间 $t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n$, 不重叠的增量

$$W_{t_2} - W_{t_1},\ W_{t_3} - W_{t_2}, \cdots, W_{t_n} - W_{t_{n-1}}$$

是相互独立的随机变量.

带漂移的布朗运动

也可以定义带漂移的布朗运动, 就是具有 $\{\mu t + W_t | t \geqslant 0\}$ 这种形式的随机过程, 其中 μ 是常数, $\{W_t\}$ 是布朗运动. 下面为正式的定义.

定义 2 一个连续随机过程 $\{W_t|\ t \geqslant 0\}$ 是波动率为 σ 漂移系数为 μ 的布朗运动或者维纳过程, 如果

1) $W_0 = 0$.

2) W_t 服从均值为 μt 方差为 $\sigma^2 t$ 的正态分布.

3) 过程 $\{W_t\}$ 具有平稳增量性质, 因此 $W_t - W_s$ 服从均值为 $\mu(t-s)$ 方差为 $\sigma^2(t-s)$ 的正态分布.

4) $\{W_t\}$ 具有独立增量性质.

样本路径

图 1 给出了漂移系数为 $\mu = 0.08$ 波动率为 $\sigma = 0.20$ 的布朗运动在区间 $[0,1]$ 上的模拟路径. 直线给出了漂移量.

更具体地, 如果固定一个结果 $\omega \in \Omega$, 那么就能定义一个函数

$$t \to W_t(\omega).$$

该函数的图像就被称为一条样本路径.

图 2 给出一个布朗运动 (与上面的布朗运动一样, 只是除了 $\sigma = 0.05$ 外) 的离散样本路径. 如所看到的, 波动率可以说是名副其实.

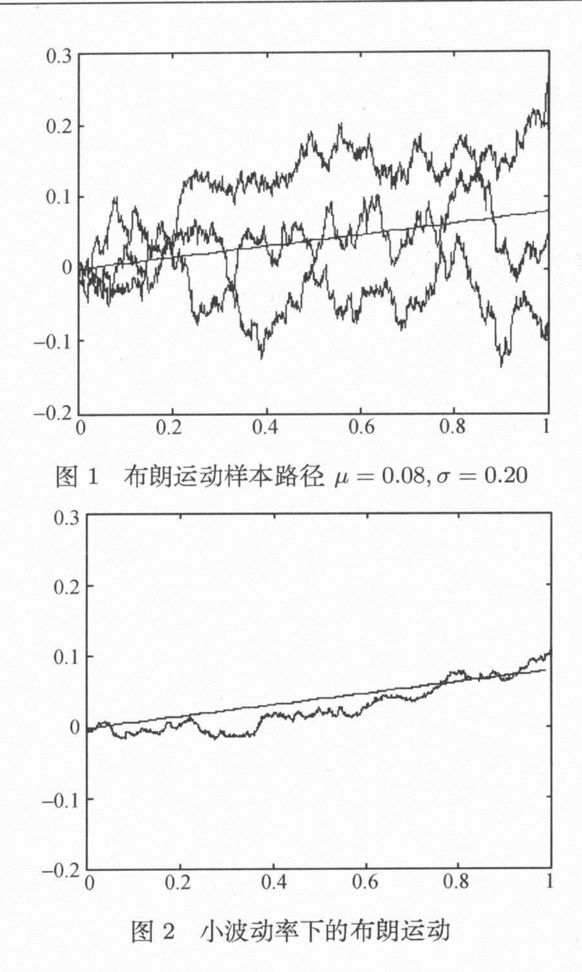


图 1　布朗运动样本路径 $\mu = 0.08, \sigma = 0.20$

图 2　小波动率下的布朗运动

布朗运动是随机过程中最重要的一类, 应用非常广泛. 我们不应该对此感到惊讶, 如果从中心极限定理的角度来看. 中心极限定理告诉我们正态分布是如何的重要 (布朗运动是一种“滑动”的正态分布).

不管如何, 布朗运动有一些非常特殊的性质. 例如, 样本路径总是连续函数. 换句话说, 样本路径没有任何跳跃点. 另一方面, 所有的路径也都是几乎处处不可微的. 也就是说, 不能在曲线上的任何位置定义切线. 因此, 样本路径没有跳跃但是变化非常剧烈, 连续突然地改变方向 (前面的图不能判断该结论, 因为它们都不是真实的样本路径).

标准布朗运动

称一个布朗运动 $\{W_t|\ t \geqslant 0\}$ 是标准布朗运动, 如果 $\mu = 0, \sigma = 1$. 在该情形下, W_t 的均值为 0, 方差为 t. 如果 $\{W_t|\ t \geqslant 0\}$ 是漂移系数为 μ 方差为 σ^2 的布朗运动,

我们能将它写成

$$W_t = \mu t + \sigma Z_t,$$

其中 $\{Z_t|\ t \geqslant 0\}$ 是标准布朗运动.

几何布朗运动与股票价格

带漂移的布朗运动是如何与股票价格联系起来的? 一种想法就是将股票价格看作“微粒”, 它受到“更小微粒”, 如股票交易或者其他局部事件, 持续的强烈撞击. 更能支持该观点的是, 我们注意到正态分布看起来就是一个用来刻画股票价格的合理选择, 如果股价的变化看作是因大量或多或少的独立 (和相似分布) 因素引起的.

然而, 用布朗运动的观点来看股票价格其本身存在明显的问题. 首先, 布朗运动可能为负而股价永远不会是负的. 其次, 在布朗运动中, 增量

$$W_t - W_s$$

的分布仅依赖 $t-s$. 因此, 如果股票的价格表现像一个布朗运动 W_t, 那么在一段时间内股价的期望变化 $\mathcal{E}(W_t - W_s)$ 就应该是 $\mu(t-s)$, 这不依赖初始价格 W_s. 这是不现实的. 例如, 想象一下时间间隔长度为 $t-s$, 价格变化为 $\mu(t-s) = \$10$. 如果股票的初始价格为 $W_s = \$100$, \$10 的期望价格变化应该是很合理的, 但是如果股票的初始价格为 $W_s = \$1$, 这就不那么合理了.

为了解决这些问题, 将股票价格的收益率看作布朗运动似乎比较合理些, 因为从数量上看, 改变量独立于初始价格似乎更合理. 例如, 说股票的收益率是 10%, 就是说股价可能从 \$100 上涨到 \$110 或者从 \$1 上涨到 \$1.10.

这可通过假设股票在时间 t 的价格 S_t 的表达式来实现, S_t 由下式给出

$$S_t = S_0 \mathrm{e}^{H_t},$$

其中 S_0 是初始价格, H_t 是布朗运动. 指数 H_t 代表股票在时间 $[0,t]$ 内的连续复合收益率. 也可以将 H_t 看作是股价的对数增长率, 因为注意到 H_t 满足

$$H_t = \log\left(\frac{S_t}{S_0}\right).$$

定义 3 *称形如 $\{\mathrm{e}^{W_t}|\ t \geqslant 0\}$ 的随机过程为几何布朗运动, 这里 $\{W_t|\ t \geqslant 0\}$ 是一个布朗运动.*

图 3 给出了几何布朗运动的一条模拟样本路径, 其中布朗运动的漂移系数与先前的一样 ($\mu = 0.08$) 但波动率超乎寻常的大, 这是为了示范增长的指数性质.

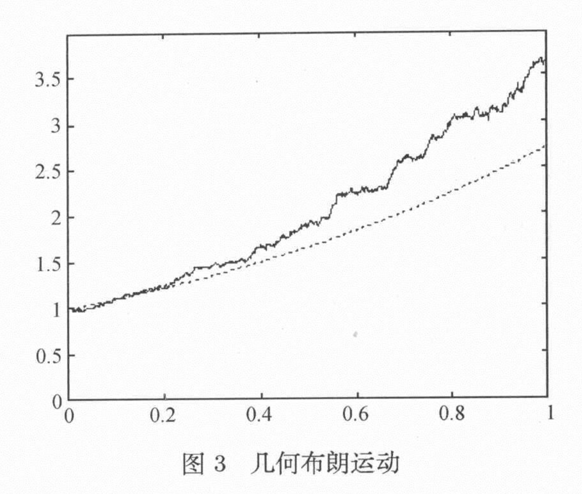


图 3　几何布朗运动

如果假定 H_t 是一个带漂移的布朗运动, 那么可以写成

$$H_t = \log\left(\frac{S_t}{S_0}\right) = \mu t + \sigma W_t,$$

其中 $\{W_t\}$ 是标准布朗运动. 因此 H_t 服从正态分布, 均值和方差为

$$\mathcal{E}(H_t) = \mu t,$$
$$\mathrm{var}(H_t) = \sigma^2 t.$$

如我们所看到的, 如果一个随机变量 X 具有性质：它的对数 $\log X$ 服从正态分布, 那么我们称随机变量 X 服从对数正态分布.

因而, S_t/S_0 服从对数正态分布. 图 4 画出了一个典型的对数正态密度函数.

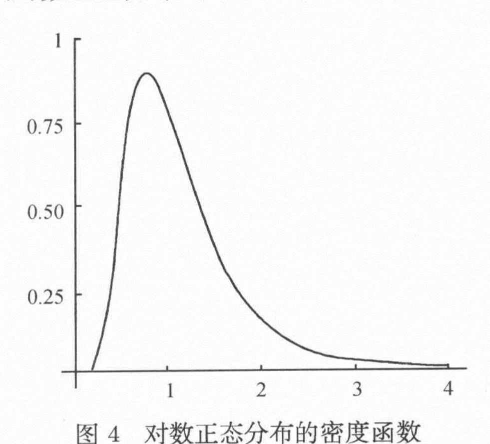


图 4　对数正态分布的密度函数

根据第 8 章的定理 9, 有

$$\mathcal{E}(S_t) = S_0 e^{(\mu+\frac{1}{2}\sigma^2)t},$$
$$\operatorname{var}(S_t) = (S_0 e^{(\mu+\frac{1}{2}\sigma^2)t})^2(e^{\sigma^2 t} - 1),$$

这里 $\mathcal{E}(S_t)$ 的值十分有趣, 因为它告诉我们期望股价不仅依赖 H_t 的漂移系数 μ, 而且还依赖波动率 σ. 从数学上来看, 我们不应该对此感到吃惊, 因为没有规则说一个随机变量的函数的均值仅仅是关于随机变量的均值的函数.

模型的另一种求解方法

如先前提到的那样, 求解连续模型的方法就是将它看作 CRR 模型的极限情况. 我们尽量在该方法上做到数学上的严谨. 然而, 我们需要花几分钟来讨论一下, 一般地怎样求解该模型才算是比较严谨的. 因为正式的讨论要求相当多的数学知识, 这超过我们现在所知道的, 因此我们非正式地来处理它.

首先, 进一步地来看看股票收益率这个概念. 该术语不只一种意思. 我们已经考虑了 H_t 在时间 $[0,t]$ 内的连续复合收益率, 该收益率满足方程

$$S_t = S_0 e^{H_t}.$$

股价在短期 $[t, t+\Delta t]$ 内的简单收益率由下式给出

$$\frac{\Delta S_t}{S_t} = \frac{S_{t+\Delta t} - S_t}{S_t},$$

当 $\Delta t \to 0$ 时, 认为上述简单收益率是无穷小时间间隔 $\mathrm{d}t$ 上的收益率. 称它为股价的瞬间百分比收益 (instantaneous percentage return) 更合适, 记为

$$\frac{\mathrm{d}S_t}{S_t}.$$

现在, 求解股票价格连续模型的共同方法就是假设瞬间百分比收益是一个布朗运动过程, 具体就是

$$\frac{\mathrm{d}S_t}{S_t} = \mu_0 \mathrm{d}t + \sigma_0 \mathrm{d}W_t,$$

其中 $\{W_t \mid t \geqslant 0\}$ 是一个标准布朗运动, μ_0 和 σ_0 是常数. 对我们来说该方程仍然有点含糊, 因为微分项 $\mathrm{d}S_t$ 和 $\mathrm{d}W_t$ 的含义特别棘手. 但是, 可以假定随机过程

$$\left\{ \frac{\mathrm{d}S_t}{S_t} \,\middle|\, t \geqslant 0 \right\}$$

(不管实际上是什么) 是一个带漂移 μ_0 波动率为 σ_0 的布朗运动.

前面的公式可以写成

$$\mathrm{d}S_t = \mu_0 S_t \mathrm{d}t + \sigma_0 S_t W_t.$$

这就是常说的随机微分方程的一个例子, 求解随机微分方程需要一些更深奥的数学知识, 即随机积分. 非常幸运的是, 我们的目的是通过 CRR 模型得到一个关于股价本身的表达式, 而不需要求解上述微分方程.

9.2　CRR 模型的极限：布朗运动

将模型的时间具体化

$$t_0 = 0 < t_1 < \cdots < t_{T-1} < t_T = t.$$

因此, 模型的整个周期为 $[0, t]$. 注意, 我们用 t 来代替 L 是因为想将 t 看作一个变量. T 个区间中的每个区间长度均为

$$t_{k+1} - t_k = \Delta t = \frac{t}{T}.$$

在每个区间内, 股价可能上涨到原来的 u 倍也可能下跌至 d 倍. 模型的样本空间就是状态空间

$$\Omega_T = \{U, D\}^T,$$

这里 $\{U, D\}^T$ 表示由 U 和 D 组成的长度为 T 的所有序列.

我们将涉及 Ω_T 上的两个不同概率测度：客观概率测度和鞅测度. 记股价上涨的客观概率为 ν, 鞅测度对应的上涨概率为 π(丢掉了前几章中使用的下标 U). 在处理这两种概率以及它们的相互作用之前, 用一个任意的概率值 p 来表示股价上涨的概率, 然后在此之上来看对数价格的增长, 这对澄清后面的说明有帮助 ($q = 1 - p$ 为下跌的概率).

令 E_i 给出了股票在区间 $[t_i, t_{i+1}]$ 上的价格运动, 即对任意终端状态 $\omega = e_1 \cdots e_T \in \Omega_T (e_i = U \text{或者} D)$, 令

$$E_i(\omega) = \begin{cases} u, & \text{如果 } e_i = U(\text{在时刻 } t_i \text{上涨}), \\ d, & \text{如果 } e_i = D(\text{在时刻 } t_i \text{下跌}). \end{cases}$$

故股票在最后时刻 t 的价格为

$$S_T = S_0 E_1 \cdots E_T = S_0 \mathrm{e}^{\sum \log(E_i)} = S_0 \mathrm{e}^{H_T},$$

其中

$$H_T = \sum_{i=1}^{T} \log(E_i) = \log\left(\frac{S_T}{S_0}\right)$$

为股价的对数增长率. 如果记股票上涨的概率为 p, 那么可定义 μ_p 和 σ_p 如下

$$\mu_p = \frac{1}{\Delta t}\mathcal{E}_p(\log E_i) = \frac{1}{\Delta t}(p\log u + q\log d),$$
$$\sigma_p^2 = \frac{1}{\Delta t}\mathrm{var}_p(\log E_i) = \frac{1}{\Delta t}pq(\log u - \log d)^2.$$

因此, 每单位时间内对数价格变化量 $\log(E_i)$ 的期望为 μ_p, 方差为 σ_p^2.

又因为有

$$\mathcal{E}_p(\log E_i) = \mu_p \Delta t$$

和

$$\mathrm{var}_p(\log E_i) = \sigma_p^2 \Delta t.$$

对随机变量 $\log E_i$ 进行标准化, 得 (因为 $\sigma \neq 0$)

$$X_{p,i} = \frac{\log E_i - \mu_p \Delta t}{\sigma_p \sqrt{\Delta t}}.$$

现在将 $\log E_i$ 写成如下形式

$$\log E_i = \mu_p \Delta t + \sigma_p \sqrt{\Delta t}\left[\frac{\log E_i - \mu_p \Delta t}{\sigma_p \sqrt{\Delta t}}\right] = \mu_p \Delta t + \sigma_p \sqrt{\Delta t} X_{p,i},$$

其中

$$X_{p,i} = \frac{\log E_i - \mu_p \Delta t}{\sigma_p \sqrt{\Delta t}}.$$

通过运用一点点代数知识处理后, 有

$$X_{p,i} = \begin{cases} \dfrac{q}{\sqrt{pq}}, & \text{如果 } e_i = U, \\ \dfrac{-p}{\sqrt{pq}}, & \text{如果 } e_i = D. \end{cases}$$

对数增长率为

$$\begin{aligned} H_T &= \sum_{i=1}^{T} \log(E_i) \\ &= \sum_{i=1}^{T} [\mu_p \Delta t + \sigma_p \sqrt{\Delta t} X_{p,i}] \end{aligned}$$

$$= \mu_p t + \sigma_p \sqrt{\Delta t} \sum_{i=1}^{T} X_{p,i}.$$

注意到公式

$$H_T = \mu_p t + \sigma_p \sqrt{\Delta t} \sum_{i=1}^{T} X_{p,i}$$

对任何 $0 < p < 1$ 都是正确的. 称 μ_p 为股价的漂移系数, σ_p 为股价的波动率. 这些量是相对于上涨概率为 p 来说的, 正如下标所暗示的那样. 但是, 注意到 H_T 本身不依赖 p. 它仅仅是状态空间 Ω_T 上的一个函数, 对每个 $p \in (0,1)$ 有不同的表达形式而已.

关于概率方面的更多知识

如前面提到的, p 为股票价格在一个子区间内上涨的概率, 然后取 $p = \nu$ 或 $p = \pi$, 但是我们现在不想做这个限制了.

利用概率值 p 可以在状态空间 Ω_T 上定义一个概率测度 $\mathbb{P}_p$ 如下

$$\mathbb{P}_p(\omega) = p^{N_U(\omega)} q^{\operatorname{len}(\omega) - N_U(\omega)},$$

$\operatorname{len}(\omega)$ 表示 ω 的长度. 由 $X_{p,i}$ 的定义可知

$$\mathbb{P}_p\left(X_{p,i} = \frac{q}{\sqrt{pq}}\right) = \mathbb{P}(e_i = U) = p,$$

$$\mathbb{P}_p\left(X_{p,i} = \frac{-p}{\sqrt{pq}}\right) = \mathbb{P}(e_i = D) = q.$$

所以随机变量 $X_{p,i}$ 在 $\mathbb{P}_p$ 下服从伯努利分布.

也能计算 $X_{p,i}$ 在 $\mathbb{P}_p$ 下的期望和方差

$$\mathcal{E}_p(X_{p,i}) = \frac{q}{\sqrt{pq}} p - \frac{p}{\sqrt{pq}} q = 0$$

和

$$\operatorname{var}_p(X_{p,i}) = 1.$$

将验证过程留作练习.

现在可将对数增长率 H_T 的表达式

$$H_T = \mu_p t + \sigma_p \sqrt{\Delta t} \sum_{i=1}^{T} X_{p,i}$$

看成两部分之和：确定部分 $\mu_p t$, 为一个常数乘以时间, 因此它像无风险资产一样以固定比率 μ_p 增长, 和随机部分

$$Q_p = \sigma_p \sqrt{\Delta t} \sum_{i=1}^{T} X_{p,i},$$

该部分为 $\sigma_p \sqrt{\Delta t}$ 乘以伯努利随机变量的和. 因为求和中的各项是独立的, 所以有

$$\mathcal{E}_p(Q_p) = 0,$$
$$\mathrm{var}_p(Q_p) = \sigma_p^2 t.$$

9.3 $\Delta t \to 0$ 时的极限

取 $T \to \infty$ 或等价地 $\Delta t \to 0$ 时的极限, 我们想仔细地弄明白哪些量会随着 T 变化. 也想弄明白哪些量依赖周期时间 t, 在这里将 t 看作一个变量. 因此改变符号如下：

- 记 $S_{t,T}$ 为最终的股价, $H_{t,T}$ 为对数增长率. 初始股价 S_0 不依赖 T, 故我们不必改变这个符号.
- 记 u_T 和 d_T 为股价上涨和下跌的因子系数.
- 记 p_T 为股价上涨的概率.
- 记 $\mu_{p_T,T}$ 和 $\sigma_{p_T,T}$ 为漂移系数和波动率.
- 记 $X_{p_T,T,i}$ 表示随机变量 $X_{p_T,i}$.

利用上面这些符号, H_t 的表达式就变成了

$$H_{t,T} = \mu_{p_T,T} t + \sigma_{p_T,T} \sqrt{\Delta t} \sum_{i=1}^{T} X_{p_T,T,i},$$

其确定部分为 $\mu_{p_T,T} t$, 随机部分为

$$Q_{p_T,T} = \sigma_{p_T,T} \sqrt{\Delta t} \sum_{i=1}^{T} X_{p_T,T,i},$$

对 $Q_{p_T,T}$ 有

$$\mathcal{E}_p(Q_{p_T,T}) = 0,$$
$$\mathrm{var}_p(Q_{p_T,T}) = \sigma_{p_T,T}^2 t.$$

现在我们想对随机部分运用中心极限定理. 为了这个目的, 考虑由伯努利随机变量形成的三角阵

$$
\begin{array}{llll}
X_{p_1,1,1} & & & \\
X_{p_2,2,1} & X_{p_2,2,2} & & \\
X_{p_3,3,1} & X_{p_3,3,2} & X_{p_3,3,3} & \\
\vdots & \vdots & \vdots & \ddots
\end{array}
$$

对每个固定的 T, 即排列中的每一行, 这些随机变量相互独立 (由 CRR 模型的假设) 且满足

$$
\mathbb{P}_{p_T}\left(X_{p_T,T,i} = \frac{1-p_T}{\sqrt{p_T(1-p_T)}}\right) = p_T,
$$
$$
\mathbb{P}_{p_T}\left(X_{p_T,T,i} = \frac{-p_T}{\sqrt{p_T(1-p_T)}}\right) = 1-p_T.
$$

因此它们也是同分布的. 但是要注意, $X_{p_T,T,i}$ 是定义在概率空间 $(\Omega_T, \mathbb{P}_{p_T})$ 上的随机变量, 因此对不同的 T(即排列中不同的行), 随机变量 $X_{p_T,T,i}$ 定义在不同的概率空间上. 这就是我们为什么需要有第 8 章定理 13 中那种形式的中心极限定理, 假如概率值 p_T 满足下面的条件, 那么就可应用该定理了.

必要条件 1: $p_T \to p$, 对某个 $p \in (0,1)$.

假设该条件满足, 我们能推出随机变量

$$
Y_{p_T,T} = \frac{1}{\sqrt{T}}\sum_{i=1}^{T} X_{p_T,T,i} = \frac{1}{\sqrt{t}}\sqrt{\Delta t}\sum_{i=1}^{T} X_{p_T,T,i} = \frac{1}{\sigma_{p_T,T}\sqrt{t}} Q_{p_T,T}
$$

依分布收敛到标准正态分布, 即

$$
Q^*_{p_T,T} = \frac{Q_{p_T,T} - \mathcal{E}_{p_T}(Q_{p_T,T})}{\sqrt{\operatorname{var}_{p_T}(Q_{p_T,T})}} = \frac{Q_{p_T,T}}{\sigma_{p_T,T}\sqrt{t}} \xrightarrow{\text{dist}} Z_t,
$$

其中 Z_t 为某个概率空间上的任一标准正态随机变量. 为了更明确些, 对任何实数 s, 有

$$
\lim_{T\to\infty} \mathbb{P}_{p_T}\left(\frac{Q_{p_T,T}}{\sigma_{p_T,T}\sqrt{t}} < s\right) = \phi_{0,1}(s),
$$

其中 $\phi_{0,1}$ 是标准正态分布函数. 为了强调收敛是在概率测度 $\mathbb{P}_{p_T}$ 下的, 我们写成

$$
\frac{Q_{p_T,T}}{\sigma_{p_T,T}\sqrt{t}} \xrightarrow{\text{dist}(p_T)} Z_t.
$$

我们现在想推出 $H_{t,T}$ 本身也依分布收敛到某一随机变量. 为了达到目的, 我们来回顾一下第 8 章的定理 11, 该定理说如果满足

必要条件 2: $\mu_{p_T,T} \to \mu, \sigma_{p_T,T} \to \sigma$, 对某个实数 μ 和 $\sigma \neq 0$.

那么有如下的极限

$$H_{t,T} = \mu_{p_T,T}t + \sigma_{p_T,T}\sqrt{t}\left(\frac{1}{\sigma_{p_T,T}\sqrt{t}}Q_{p_T,T}\right) \xrightarrow{\text{dist}(p_T)} \mu t + \sigma\sqrt{t}Z_t.$$

关于股票价格本身, 因为 $S_{t,T}$ 是 $H_{t,T}$ 的连续函数

$$S_{t,T} = S_0\mathrm{e}^{H_{t,T}}.$$

第 8 章定理 10 告诉我们, 只要满足必要条件 1 和 2, 就有

$$S_{t,T} \xrightarrow{\text{dist}(p_T)} S_0\mathrm{e}^{\mu t+\sigma\sqrt{t}Z_t},$$

这里 Z_t 为标准正态分布. 令

$$W_t = \sqrt{t}Z_t,$$

得到

$$H_{t,T} \xrightarrow{\text{dist}(p_T)} H_t = \mu t + \sigma W_t$$

和

$$S_{t,T} \xrightarrow{\text{dist}(p_T)} S_t = S_0\mathrm{e}^{\mu t+\sigma W_t},$$

这里

$$\mathcal{E}(W_t) = 0,$$
$$\text{var}(W_t) = t.$$

布朗运动和几何布朗运动

虽然前面已经推出了周期 t 固定时的一些公式, 但是我们想考虑 t 为变量的情形. 非常遗憾的是, 当 t 取遍所有的非负实数时, 我们的推导不能直接揭露出不同随机变量 W_t 之间的重要关系.

我们就不正式地探究该问题, 但是能做一点非正式的观察. 首先, 从直觉上应该清楚地知道我们的模型是“交换不变的”或“稳定的”, 意思是在区间 $[s,t]$ 和 $[0,t-s]$ 上我们得到的模型在本质上是一样的. 其次, 因为假定各子区间上的变化是

独立的, 所以应该能够把不相交的临近区间的模型连接起来成为一个大区间上的模型. 因此, 当 t 变化时, 我们不应该对如下的随机过程是一个标准布朗运动而感到奇怪,

$$\{W_t|\ t \geqslant 0\}.$$

故

$$\{H_t|\ t \geqslant 0\}$$

是均值为 μt 方差为 $\sigma^2 t$ 的布朗运动, 即漂移系数为 μ 及波动率为 σ. 最后, 股票价格过程本身

$$\{S_t|\ t \geqslant 0\}$$

是一个几何布朗运动, 其漂移系数及波动率分别为 μ 和 σ.

现在来总结一下学过的知识, 即股价在 CRR 模型的极限情形下的表现行为.

定理 1　设 CRR 模型中的上涨概率为 p_T, 模型的整个时间区间为 $[0,t]$, T 为等长度 Δt 的子区间个数, 令 $S_{t,T}$ 为 CRR 模型下的最终股价. 且假定

1) $p_T \to p$, 对某个 $p \in (0,1)$,

2) $\mu_{p_T,T} \to \mu, \sigma_{p_T,T} \to \sigma$, 对某个实数 μ 和 $\sigma \neq 0$.

令

$$H_{t,T} = \log\left(\frac{S_{t,T}}{S_0}\right)$$

为价格对数增长率. 那么有

$$H_{t,T} \xrightarrow{\text{dist}(p_T)} H_t = \mu t + \sigma W_t$$

和

$$S_{t,T} \xrightarrow{\text{dist}(p_T)} S_t = S_0 \mathrm{e}^{\mu t + \sigma W_t},$$

其中

$$\{W_t|\ t \geqslant 0\}$$

是标准布朗运动. 因此, 对数增长率

$$\{H_t|\ t \geqslant 0\}$$

是均值为 μt 方差为 $\sigma^2 t$ 的布朗运动, 即漂移系数为 μ 和波动率为 σ. 股票价格过程本身

$$\{S_t|\ t \geqslant 0\}$$

是一个几何布朗运动, 其漂移系数和波动率分别为 μ 和 σ. 也注意到股价增长率 S_t/S_0 是对数正态分布的, 其均值和方差为

$$\mathcal{E}(S_t) = S_0 \mathrm{e}^{(\mu+\frac{1}{2}\sigma^2)t},$$
$$\mathrm{var}(S_t) = (S_0 \mathrm{e}^{(\mu+\frac{1}{2}\sigma^2)t})^2(\mathrm{e}^{\sigma^2 t} - 1).$$

9.4 客观概率下的 CRR 模型

对一般的 CRR 模型来说, 我们已经讨论了能够讨论的, 现在来思考如何构造一个模型能反映“真实”世界. 称下面的模型为客观概率下的 CRR 模型, 记为 CRR_ν.

首先, 假定存在一个概率测度, 它反映了市场上涨的真实情况, 称其为客观概率. 习惯对客观概率作如下假设.

假设 1 在模型整个周期 T 内, 上涨的客观概率为常数. 将该概率值记为 ν.

也习惯作出下面的假设.

假设 2 在客观概率下, 漂移系数和波动率关于 Δt(和 T) 为常数, 即

$$\mu_\nu = \frac{1}{\Delta t}(\nu \log u_T + (1-\nu)\log d_T),$$

$$\sigma_\nu^2 = \frac{1}{\Delta t}\nu(1-\nu)(\log u_T - \log d_T)^2.$$

因此, 我们可以省略下标 T 而写成 μ_ν 和 σ_ν. 称 μ_ν 为瞬时漂移系数, σ_ν 为瞬时波动率.

这里需要强调的是, 第二个假设在模型的上涨因子、下跌因子和鞅测度方面有一些重要的结果.

在一般的 CRR 模型中, u_T, d_T 和 $p_T(=\nu)$ 是不相关的, 尽管漂移系数和波动率是用这些量来定义的. 然而, 具体化 μ_ν 和 σ_ν 这些特定的常数等于具体化 u_T, d_T 和 ν 之间的关系. 为了描绘类似的简单关系, 假设 x 和 y 是任意的随机变量, 定义 A 如下

$$A = x + y.$$

比如一旦假定 $A = 5$, 我们就描绘出了 x 和 y 之间的一种关系.

u_T, d_T 和 ν 之间的关系可简单地通过解漂移系数和波动率的定义方程得到, 有

$$\log u_{\nu,T} = \mu_\nu \Delta t + \sigma_\nu \sqrt{\Delta t} \frac{1-\nu}{\sqrt{\nu(1-\nu)}},$$
$$\log d_{\nu,T} = \mu_\nu \Delta t - \sigma_\nu \sqrt{\Delta t} \frac{\nu}{\sqrt{\nu(1-\nu)}},$$

右边部分通过 $\Delta t = t/T$ 而依赖 T. 同样, 我们写成 $u_{\nu,T}$ 和 $d_{\nu,T}$ 是为了强调它们也依赖 ν. 注意这种依赖是通过下面的鞅测度反映出来的

$$\pi_{\nu,T} = \frac{\mathrm{e}^{rT} - d_{\nu,T}}{u_{\nu,T} - d_{\nu,T}}.$$

事实上, 可以直接用 ν、漂移系数和波动率表示鞅测度. 这也导出了一个有趣的极限.

定理 2　1) 在模型 CRR_ν 中, 鞅测度的上涨概率为

$$\pi_{\nu,T} = \frac{\mathrm{e}^{(r-\mu_\nu)\Delta t + \frac{\nu}{\sqrt{\nu(1-\nu)}}\sigma_\nu\sqrt{\Delta t}} - 1}{\mathrm{e}^{\frac{1}{\sqrt{\nu(1-\nu)}}\sigma_\nu\sqrt{\Delta t}} - 1}.$$

2) 当 $T \to \infty$ 或等价地 $\Delta t \to \infty$ 时, 鞅概率 $\pi_{\nu,T}$ 趋于客观概率 ν, 即

$$\lim_{T\to\infty} \pi_{\nu,T} = \nu.$$

证明　1) 鞅测度的上涨概率由下式给出

$$\pi_{\nu,T} = \frac{\mathrm{e}^{rT} - d_{\nu,T}}{u_{\nu,T} - d_{\nu,T}} = \frac{\mathrm{e}^{r\Delta t}(d_{\nu,T})^{-1} - 1}{u_{\nu,T}(d_{\nu,T})^{-1} - 1}.$$

由前面关于 $\log u_{\nu,T}$ 和 $\log d_{\nu,T}$ 的方程可得到

$$u_{\nu,T} = \mathrm{e}^{\mu_\nu \Delta t + \sigma_\nu \sqrt{\Delta t}\frac{1-\nu}{\sqrt{\nu(1-\nu)}}},$$
$$(d_{\nu,T})^{-1} = \mathrm{e}^{-\mu_\nu \Delta t + \sigma_\nu \sqrt{\Delta t}\frac{\nu}{\sqrt{\nu(1-\nu)}}}.$$

所以有

$$\mathrm{e}^{r\Delta t}(d_{\nu,T})^{-1} = \mathrm{e}^{(r-\mu_\nu)\Delta t + \sigma_\nu\sqrt{\Delta t}\frac{\nu}{\sqrt{\nu(1-\nu)}}}$$

和

$$u_{\nu,T}(d_{\nu,T})^{-1} = \mathrm{e}^{\sigma_\nu\sqrt{\Delta t}\frac{1}{\sqrt{\nu(1-\nu)}}}.$$

将上述两个式子代入前面 $\pi_{\nu,T}$ 的表达式, 立即得到要证的公式.

2) 运用洛必达法则就可以求出极限

$$\lim_{T\to\infty}\pi_{\nu,T}=\lim_{\Delta t\to\infty}\frac{\mathrm{e}^{(r-\mu_\nu)\Delta t+\frac{\nu}{\sqrt{\nu(1-\nu)}}\sigma_\nu\sqrt{\Delta t}}-1}{\mathrm{e}^{\frac{1}{\sqrt{\nu(1-\nu)}}\sigma_\nu\sqrt{\Delta t}}-1}.$$

我们将过程留给读者.

瞬时漂移系数和瞬时波动率为常数这个假设可能存在一点问题 (与假设客观概率为常数一样), 但是这样做是基于一些实际考虑 (就像经常有疑问的假设一样). 事实上, 经常将该假设扩展到过去. 特别地, 假定在客观概率下可通过股价的历史价格来计算 (至少可估计) 漂移系数和波动率, 这也是值得怀疑的.

具体地, 在客观概率下估计瞬时漂移系数和瞬时波动率的经验步骤如下. 首先我们为 Δt 选一个很小的值 (例如, Δt 可以对应一天). 然后从这些短期时间段内的大量数据出发来计算对数增长因子

$$\log E_i=\log\left(\frac{\text{单期的期末股价}}{\text{单期的期初股价}}\right).$$

这些样本值的平均值就是 $\mu_\nu\Delta t$ 的一个估计, 样本值的样本方差是 $\sigma_\nu^2\Delta t$ 的一个估计. 当然, 我们取的样本越多, 得到的估计就会越好. 下面来看一个例子.

例 1 下面为一张 Excel 表格的一部分, 给出了最近 10 天的股价. 这里取 Δt 为 1 天, 即 1/365 年. 初始价格为 \$50.

天数	价格	增长率	对数增长率	平均值	样本方差
0	50			0.000359354	0.000677264
1	50.95	1.019	0.018821754	Per Unit	Per Unit
2	49.74	0.976251	−0.024035321	0.131164046	0.247201227
3	49.46	0.994371	−0.005645176		
4	49.83	1.007481	0.00745295		
5	48.7	0.977323	−0.022938182		
6	50.2	1.030801	0.030335997		
7	49.57	0.98745	−0.012629215		
8	51.78	1.044583	0.043618162		
9	52.17	1.007532	0.007503643		
10	50.18	0.961855	−0.038891076		

增长率这一列表示现在股价与前一期股价的比. 例如

$$\frac{50.95}{50}=1.019.$$

平均值就是简单地将对数增长率的和除以对数增长率的数目. 样本方差由下面的

公式计算得到

$$样本方差 = \frac{1}{n-1}\sum_{i=1}^{n}(第\ i\ 个值 - 平均值)^2,$$

这里除以 $n-1$ 而不是 n, 是为了得到无偏估计. 最后, “每个单位” 值通过平均值和方差乘以 $1/\Delta t = 365$ 得到. 故以年计算的 $\mu_\nu \approx 0.13, \sigma_\nu^2 \approx 0.25$ (也就是说, 时间的单位是以年来表示的).

利用前面提到的有关客观概率的假设, 定理 1 的条件

1) $p_T \to p$, 对某个 $p \in (0,1)$,

2) $\mu_{p_T,T} \to \mu, \sigma_{p_T,T} \to \sigma$, 对某个实数 μ 和 $\sigma \neq 0$

就变得很显然了. 因为对所有的 T, 有 $p_T = \nu$, 所以有 $p = \nu$. 又因为对所有的 T 有 $\mu = \mu_\nu$ 和 $\sigma_{p_T,T} = \sigma_\nu$, 故有 $\mu = \mu_\nu$ 和 $\sigma = \sigma_\nu$. 因此, 在客观概率下定理 1 的条件成立, 我们有下面的结论.

定理 3 设 $S_{t,T}$ 是模型 CRR_ν 下的最终股价. 令

$$H_{t,T} = \log\left(\frac{S_{t,T}}{S_0}\right)$$

为对数增长率. 那么有

$$H_{t,T} \xrightarrow{\mathrm{dist}(\nu)} H_t = \mu_\nu t + \sigma_\nu W_{\nu,t}$$

和

$$S_{t,T} \xrightarrow{\mathrm{dist}(\nu)} S_t = S_0 \mathrm{e}^{\mu_\nu t + \sigma_\nu W_{\nu,t}},$$

其中

$$\{W_{\nu,t} |\ t \geqslant 0\}$$

是标准布朗运动. 因此, 对数增长率

$$\{H_t |\ t \geqslant 0\}$$

是均值为 $\mu_\nu t$ 方差为 $\sigma_\nu^2 t$ 的布朗运动, 其漂移系数为 μ_ν, 波动率为 σ_ν. 股票自身价格过程

$$\{S_t |\ t \geqslant 0\}$$

是一个几何布朗运动, 其漂移系数和波动率也分别为 μ_ν 和 σ_ν. 也注意到股价增长率 S_t/S_0 是对数正态分布的, 从而有

$$\begin{aligned}
\mathcal{E}(S_t) &= S_0 \mathrm{e}^{(\mu_\nu + \frac{1}{2}\sigma_\nu^2)t},\\
\mathrm{var}(S_t) &= (S_0 \mathrm{e}^{(\mu_\nu + \frac{1}{2}\sigma_\nu^2)t})^2(\mathrm{e}^{\sigma_\nu^2 t} - 1).
\end{aligned}$$

9.5 等价鞅测度下的 CRR 模型

先来看我们的主要目标,好让自己知道下一步要做什么. (例如) 执行价为 K 的欧式卖权的收益是

$$\begin{aligned}X &= \max\{K - S_{t,T}, 0\}\\ &= \max\{K - S_0\mathrm{e}^{H_{t,T}}, 0\}.\end{aligned}$$

无套利意味着卖权的初始价格必须是

$$\mathcal{I}(\text{卖权}) = \mathrm{e}^{-rt}\mathcal{E}_{\Pi_T}(\max\{K - S_0\mathrm{e}^{H_{t,T}}, 0\}),$$

其中 Π_T 是鞅测度 (上涨的概率为 π_T). 当 T 趋于无穷大时, 得

$$P_\infty = \lim_{T\to\infty}\mathcal{I}(\text{卖权}) = \mathrm{e}^{-rt}\lim_{T\to\infty}\mathcal{E}_{\Pi_T}(\max\{K - S_0\mathrm{e}^{H_{t,T}}, 0\}),$$

其中 P_∞ 表示价格随机变量的极限. 令

$$g(x) = \max\{K - S_0\mathrm{e}^{x}, 0\},$$

该函数在 $\mathbb{R}$ 上有界且连续, 从而

$$P_\infty = \mathrm{e}^{-rt}\lim_{T\to\infty}\mathcal{E}_{\Pi_T}(g(H_{t,T})).$$

现在, 我们想把极限运算交换到期望算子里面去而得到

$$P_\infty = \mathrm{e}^{-rt}\mathcal{E}(g(H_t)).$$

回忆下第 8 章的定理 10, 该定理说如果

$$X_n \xrightarrow{\text{dist}(\pi_T)} X.$$

那么对所有的有界连续函数 $g:\mathbb{R}\to\mathbb{R}$, 有

$$\mathcal{E}_{\Pi_T}(g(X_n)) \to \mathcal{E}(g(X)).$$

这刚好是我们需要的, 但是为了运用该定理, 需要知道 $H_{t,T}$ 在鞅测度下是弱收敛的, 而不是像定理 3 那样在客观概率测度下.

这就告诉我们接下来该做什么. 特别地, 我们需要一个新的 CRR 模型来做下面的事情.

1) 上涨的概率应该为鞅测度的上涨概率 $\pi_{\nu,T}$, 因此, 定理 1 就可给出 $H_{t,T}$ 在鞅测度下的弱极限.

2) 同时, 模型必须保持 CRR_ν 中的“真实”股价 S_k, 在这里是通过 CRR_ν 中的 $u_{\nu,T}$ 和 $d_{\nu,T}$ 来实现的.

因此, 我们定义一个含有如下参数的新 CRR 模型.

1) 上涨概率 $u_{\nu,T}$、下跌概率 $d_{\nu,T}$ 和鞅测度的上涨概率

$$\pi_{\nu,T} = \frac{\mathrm{e}^{rT} - d_{\nu,T}}{u_{\nu,T} - d_{\nu,T}}$$

与客观概率模型 CRR_ν 里的一样. 因此股价 S_k 就是所要求的“客观”价格.

2) 上涨的概率就是鞅测度的上涨概率, 即

$$p_T = \pi_{\nu,T}.$$

因此, 漂移系数和波动率为

$$\mu_{\pi_{\nu,T},T} = \frac{1}{\Delta t}(\pi_{\nu,T} \log u_{\nu,T} + (1-\pi_{\nu,T}) \log d_{\nu,T}),$$
$$\sigma^2_{\pi_{\nu,T},T} = \frac{1}{\Delta t}\pi_{\nu,T}(1-\pi_{\nu,T})(\log u_{\nu,T} - \log d_{\nu,T})^2.$$

称该模型为鞅测度下的 CRR 模型, 记为 CRR_π.

下面的定理描述了 CRR_π 模型与 CRR_ν 模型中的漂移系数和波动率之间的相互关系.

定理 4　下面的式子成立:

$$\mu_{\pi_{\nu,T},T} = \mu_\nu + \sigma_\nu \frac{1}{\sqrt{\Delta t}} \frac{(\pi_{\nu,T} - \nu)}{\sqrt{\nu(1-\nu)}},$$
$$\sigma^2_{\pi_{\nu,T},T} = \sigma^2_\nu \frac{\pi_{\nu,T}(1-\pi_{\nu,T})}{\nu(1-\nu)}.$$

证明　为了可读性, 记 $\pi_{\nu,T} = \pi, u_{\nu,T} = u,$ 和$d_{\nu,T} = d$. 那么

$$\begin{aligned}
&\frac{\mu_{\pi_{\nu,T},T} - \mu_\nu}{\sigma_\nu} \\
&= \frac{\frac{1}{\Delta t}(\pi \log u + (1-\pi)\log d) - \frac{1}{\Delta t}(\nu \log u + (1-\nu)\log d)}{\frac{1}{\sqrt{\Delta t}}\sqrt{\nu(1-\nu)}(\log u - \log d)} \\
&= \frac{\frac{1}{\Delta t}(\pi - \nu)(\log u - \log d)}{\frac{1}{\sqrt{\Delta t}}\sqrt{\nu(1-\nu)}(\log u - \log d)} \\
&= \frac{1}{\sqrt{\Delta t}} \frac{\pi - \nu}{\sqrt{\nu(1-\nu)}}.
\end{aligned}$$

从上可以求得 $\mu_{\pi_{\nu,T},T}$, 这正是我们想要的结果. 关于 $\sigma^2_{\pi_{\nu,T},T}$ 的计算过程如下

$$\begin{aligned}\frac{\sigma^2_{\pi_{\nu,T},T}}{\sigma^2_\nu} &= \frac{\dfrac{1}{\Delta t}\pi(1-\pi)(\log u - \log d)^2}{\dfrac{1}{\Delta t}\nu(1-\nu)(\log u - \log d)^2} \\ &= \frac{\pi(1-\pi)}{\nu(1-\nu)}.\end{aligned}$$

定理得证.

为了能够在模型 CRR_π 中运用定理 1, 我们现在来计算所需的极限.

定理 5　下面的极限成立:

$$\begin{aligned}\lim_{T\to\infty} \mu_{\pi_{\nu,T},T} &= r - \frac{\sigma^2_\nu}{2}, \\ \lim_{T\to\infty} \sigma_{\pi_{\nu,T},T} &= \sigma_\nu,\end{aligned}$$

其中 r 是无风险利率.

证明　定理 4 告诉我们

$$\mu_{\pi_{\nu,T},T} = \mu_\nu + \sigma_\nu \frac{1}{\sqrt{\Delta t}} \frac{(\pi_{\nu,T} - \nu)}{\sqrt{\nu(1-\nu)}}.$$

所以

$$\lim_{T\to\infty} \mu_{\pi_{\nu,T},T} = \mu_\nu + \frac{\sigma_\nu}{\sqrt{\nu(1-\nu)}} \lim_{\Delta t\to 0} \left(\frac{\pi_{\nu,T} - \nu}{\sqrt{\Delta t}}\right).$$

现在, 可以利用定理 2 中 $\pi_{\nu,T}$ 的表达式来求极限

$$\lim_{\Delta t\to 0} \left(\frac{\mathrm{e}^{(r-\mu_\nu)\Delta t + \frac{\nu}{\sqrt{\nu(1-\nu)}}\sigma_\nu\sqrt{\Delta t}} - 1}{\mathrm{e}^{\frac{1}{\sqrt{\nu(1-\nu)}}\sigma_\nu\sqrt{\Delta t}} - 1} - \nu\right),$$

这有点繁杂. 用洛必达法则和强构造法或者符号代数软件就可得到极限为

$$\frac{2(r-\mu_\nu) - \sigma^2_\nu}{2\sigma_\nu}\sqrt{\nu(1-\nu)}.$$

所以有

$$\begin{aligned}\lim_{T\to\infty} \mu_{\pi_{\nu,T},T} &= \mu_\nu + \frac{\sigma_\nu}{\sqrt{\nu(1-\nu)}} \left[\frac{2(r-\mu_\nu) - \sigma^2_\nu}{2\sigma_\nu}\sqrt{\nu(1-\nu)}\right] \\ &= r - \frac{\sigma^2_\nu}{2}.\end{aligned}$$

这就是想要的. 关于 $\sigma_{\pi_{\nu,T},T}$ 的极限, 我们从定理 4 中的公式出发, 即

$$\sigma^2_{\pi_{\nu,T},T} = \sigma^2_\nu \left(\frac{\pi_{\nu,T}(1-\pi_{\nu,T})}{\nu(1-\nu)}\right).$$

由定理 2, 有

$$\lim_{\Delta t \to 0} \pi_{\nu,T} = \nu,$$

从而

$$\lim_{\Delta t \to 0} \frac{\pi_{\nu,T}(1-\pi_{\nu,T})}{\nu(1-\nu)} = 1.$$

故

$$\lim_{T \to \infty} \sigma^2_{\pi_{\nu,T},T} = \sigma^2_{\nu}.$$

取平方根就得到了要证的结果.

现在, 定理 1 就能给我们想要的弱极限.

定理 6　在模型 CRR_{π} 中, 鞅测度的上涨概率为 $\pi_{\nu,T}$, 时间区间为 $[0,t]$, T 等于长度为 Δt 的区间个数. 设 $S_{t,T}$ 是最终的股价. 令

$$H_{t,T} = \log\left(\frac{S_{t,T}}{S_0}\right)$$

为对数增长率. 在鞅测度下, 有

$$H_{t,T} \xrightarrow{\mathrm{dist}(\pi_T)} = H_t = \left(r - \frac{\sigma_{\nu}^2}{2}\right)t + \sigma_{\nu} W_{\pi,t}$$

和

$$S_{t,T} \xrightarrow{\mathrm{dist}(\pi_T)} = S_t = S_0 \mathrm{e}^{(r-\frac{\sigma_{\nu}^2}{2})t + \sigma_{\nu} W_{\pi,t}},$$

其中

$$\{W_{\pi,t} |\ t \geqslant 0\}$$

是一个标准布朗运动. 因此, 对数增长率

$$\{H_{\pi,t} |\ t \geqslant 0\}$$

是一个布朗运动, 且

1) 均值为 $\left(r - \dfrac{\sigma_{\nu}^2}{2}\right)t$.

2) 方差为 $\sigma_{\nu}^2 t$.

3) 漂移系数为 $r - \dfrac{\sigma_{\nu}^2}{2}$.

4) 波动率为 σ_{ν}.

股票的自身价格过程

$$\{S_t |\ t \geqslant 0\}$$

在鞅测度下是一个几何布朗运动，漂移系数和波动率与 $H_{\pi,t}$ 的一样. 也注意到股价增长率 S_t/S_0 是对数正态分布的，有

$$\mathcal{E}(S_t) = S_0 e^{rt},$$
$$\text{var}(S_t) = (S_0 e^{rt})^2 (e^{\sigma_\nu^2 t} - 1).$$

9.6 从一个不同的观点看模型：Itô 引理

在本章早些时候，我们谈到了发展连续模型的一个常用方法，即假设股票价格过程满足随机微分方程

$$\frac{dS_t}{S_t} = \mu_0 dt + \sigma_0 dW_t,$$

其中 $\{W_t | \ t \geqslant 0\}$ 是标准布朗运动，μ_0 和 σ_0 是常数 (不要与 $\mu_\nu, \mu_{\pi_T}, \sigma_\nu, \sigma_{\pi_T}$ 混淆). 同时也提到，对我们来说该方程的精确意思仍然有一点模糊，因为它要求我们有更多的数学知识，这超出了我们目前所讲的. 虽然这是真的，但是我们可以稍微“挥挥手”，然后去更深刻地理解是如何使用该方程来发展模型的. 读者阅读文献时可能遇到该方程，从这个角度出发，这至少是值得花力气的. 但是，读者也可以跳过该部分讨论而不会影响学习的连贯性.

前面的方程可以写为

$$dS_t = \mu_0 S_t dt + \sigma_0 S_t dW_t.$$

这是下面方程的特例

$$dS_t = a(S_t, t)dt + b(S_t, t)dW_t,$$

其中 $a(x,t)$ 和 $b(x,t)$ 是双变量函数. 在我们的情形中，有

$$a(S_t, t) = \mu_0 S_t,$$
$$b(S_t, t) = \sigma_0 S_t.$$

有时称满足上述方程的过程 $\{S_t\}$ 为 Itô 过程.

现在，如果 $f(x,t)$ 是一双变量函数，那么就能构成复合函数 $f(S_t, t)$，它是一个随机过程，因为 $\{S_t\}$ 是一个随机过程. 我们对找到 df 的公式感兴趣. 这可应用随机积分中一个所谓的 Itô 引理来实现.

定理 7(Itô 引理) 设 $\{S_t\}$ 满足 Itô 过程

$$dS_t = a(S_t, t)dt + b(S_t, t)dW_t,$$

其中 $\{W_t\}$ 是标准布朗运动, $a(x,t)$ 和 $b(x,t)$ 是双变量函数. 假设 $f(x,t)$ 是一个 (充分可微的) 双变量函数. 那么

$$\mathrm{d}f = \left(\frac{\partial f}{\partial x}a + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}b^2\right)\mathrm{d}t + \frac{\partial f}{\partial x}b\ \mathrm{d}W_t.$$

在我们所考虑的情形中, 有

$$\begin{aligned} a(S_t,t) &= \mu_0 S_t, \\ b(S_t,t) &= \sigma_0 S_t. \end{aligned}$$

所以 Itô 引理变为

$$\mathrm{d}f = \left(\frac{\partial f}{\partial x}\mu_0 S_t + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma_0^2 S_t^2\right)\mathrm{d}t + \frac{\partial f}{\partial x}\sigma_0 S_t\ \mathrm{d}W_t.$$

现在可以将该公式应用到如下函数中去

$$f(x,t) = \log x.$$

这时有

$$\begin{aligned} \frac{\partial f}{\partial x}(S_t,t) &= \frac{1}{S_t}, \\ \frac{\partial f}{\partial t}(S_t,t) &= 0, \\ \frac{\partial^2 f}{\partial x^2}(S_t,t) &= -\frac{1}{S_t^2}. \end{aligned}$$

从而得到

$$\begin{aligned} \mathrm{d}(\log S_t) &= \left(\frac{1}{S_t}\mu_0 S_t - \frac{1}{2}\frac{1}{S_t^2}\sigma_0^2 S_t^2\right)\mathrm{d}t + \frac{1}{S_t}\sigma_0 S_t \mathrm{d}W_t \\ &= \left(\mu_0 - \frac{1}{2}\sigma_0^2\right)\mathrm{d}t + \sigma_0\ \mathrm{d}W_t. \end{aligned}$$

这表明 $\mathrm{d}(\log S_t)$ 服从正态分布, 其均值和方差分别是 $\left(\mu_0 - \frac{1}{2}\sigma_0^2\right)\mathrm{d}t$ 和 $\sigma_0^2\ \mathrm{d}t$. 从而 $\log S_t$ 在区间 $[0,t]$ 内的改变量, 即

$$\log S_t - \log S_0$$

为相互独立的正态随机变量的和, 故也服从正态分布, 且有均值

$$\sum\left(\mu_0 - \frac{1}{2}\sigma_0^2\right)\mathrm{d}t = \left(\mu_0 - \frac{1}{2}\sigma_0^2\right)t$$

和方差

$$\sum(\sigma_0^2 \mathrm{d}t) = \sigma_0^2 t.$$

因而

$$\begin{aligned} H_t &= \log\left(\frac{S_t}{S_0}\right) = \log S_t - \log S_0 \\ &= \left(\mu_0 - \frac{1}{2}\sigma_0^2\right) t + \sigma_0 \sqrt{t} Z_t, \end{aligned}$$

这里 Z_t 是标准正态的. 因此我们看到 (一般形式下) 随机微分方程导出了与定理 6 相同的结论 (这里我们没有明确地涉及鞅测度).

9.7 假设符合实际吗

为了能继续讨论, 我们先暂停一下来评论股价服从对数正态分布, 或等价地股价的对数增长率是正态分布的这个假设是否符合现实. 在这个问题上, 已经做了很多统计工作, 结果表明增长率表现出一种所谓的尖峰态 (leptokurtosis) 现象, 这意味着:

1) 对数增长率接近均值的概率大于正态分布情形下的概率 (峰更高).

2) 对数增长率远离均值的概率大于正态分布情形下的概率 (厚尾).

另外一个统计上的证明表明正态假设可能不现实. 想要更深入地了解这个问题, 我们建议读者参考 Chriss [1997]. 当然, 这没有必要感到奇怪. 毕竟, 导出下面公式的假设都不现实,

$$H_t = \left(r - \frac{\sigma_\nu^2}{2}\right) t + \sigma_\nu W_{\pi,t}.$$

然而, 基于正态模型的布莱克 – 舒尔斯公式是应用最广泛的期权定价公式, 该公式就是我们要讨论的.

9.8 布莱克–舒尔斯期权定价公式

现在, 我们已经有了推导欧式期权的布莱克 – 舒尔斯期权定价公式所必备的工具. 再一次考虑完全的、无套利 CRR_π 模型.

执行价为 K 的欧式卖权的损益为

$$\begin{aligned} X &= \max\{K - S_{t,T}, 0\} \\ &= \max\{K - S_0 \mathrm{e}^{H_{t,T}}, 0\}. \end{aligned}$$

复制定价策略表明, 为了避免套利该卖权的初始价值必须为

$$\mathcal{I}(\text{卖权}) = \mathrm{e}^{-rt}\mathcal{E}_{\Pi_T}(\max\{K - S_0\mathrm{e}^{H_{t,T}}, 0\}),$$

其中 Π_T 是鞅测度 (上涨的概率为 π_T). 当 T 趋于无穷大时, 得

$$P_\infty = \lim_{T\to\infty}\mathcal{I}(\text{卖权}) = \mathrm{e}^{-rt}\lim_{T\to\infty}\mathcal{E}_{\Pi_T}(\max\{K - S_0\mathrm{e}^{H_{t,T}}, 0\}),$$

其中 P_∞ 表示价格随机变量的极限. 这是第 8 章定理 10 的 2) 中的一种情形, 那里函数 g 为

$$g(x) = \max\{K - S_0\mathrm{e}^x, 0\},$$

该函数在 $\mathbb{R}$ 上有界且连续. 在该情形下, 有

$$P_\infty = \mathrm{e}^{-rt}\lim_{T\to\infty}\mathcal{E}_{\Pi_T}(g(H_{t,T})).$$

又因为

$$H_{t,T} \xrightarrow{\text{dist}(\pi_T)} = H_t = \left(r - \frac{\sigma_\nu^2}{2}\right)t + \sigma_\nu\sqrt{t}Z_{\pi,t}.$$

从而由第 8 章的定理 10 可知

$$P_\infty = \mathrm{e}^{-rt}\lim_{T\to\infty}\mathcal{E}(g(H_{t,T})) = \mathrm{e}^{-rt}\mathcal{E}_{\Pi_T}(g(H_t)).$$

因为 H_t 是正态分布, 有均值

$$a = \left(r - \frac{\sigma_\nu^2}{2}\right)t$$

和方差

$$b^2 = \sigma_\nu^2 t.$$

所以有

$$P_\infty = \mathrm{e}^{-rt}\mathcal{E}(g(H_t)) = \frac{\mathrm{e}^{-rt}}{\sqrt{2\pi b^2}}\int_{-\infty}^{\infty} g(x)\mathrm{e}^{-\frac{(x-a)^2}{2b^2}}\mathrm{d}x,$$

这里平方根下面的 π 就是圆周率 pi, 而不是鞅测度中的上涨概率.

我们需要做的一切就是估计上面这个积分. 作变换

$$y = \frac{x - a}{b},$$

得到

$$P_\infty = \frac{\mathrm{e}^{-rt}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(by + a)\mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y.$$

现在, 函数

$$g(by+a)=\max\{K-S_0\mathrm{e}^{by+a},0\}$$

非零当且仅当

$$K-S_0\mathrm{e}^{by+a}>0,$$

即当且仅当

$$by+a<\log\left(\frac{K}{S_0}\right),$$

或者

$$y<\frac{1}{b}\left[\log\left(\frac{K}{S_0}\right)-a\right].$$

记右边部分为 h,

$$h=\frac{1}{b}\left[\log\left(\frac{K}{S_0}\right)-a\right].$$

我们就能将积分写成

$$P_\infty=\frac{\mathrm{e}^{-rt}}{\sqrt{2\pi}}\int_{-\infty}^{h}(K-S_0\mathrm{e}^{by+a})\mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y,$$

拆开被积项, 得

$$P_\infty=\frac{K\mathrm{e}^{-rt}}{\sqrt{2\pi}}\int_{-\infty}^{h}\mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y-\frac{S_0\mathrm{e}^{-rt}}{\sqrt{2\pi}}\int_{-\infty}^{h}\mathrm{e}^{by+a}\mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y.$$

上面积分式子的第一项是非常漂亮的, 因为有

$$\frac{K\mathrm{e}^{-rt}}{\sqrt{2\pi}}\int_{-\infty}^{h}\mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y=K\mathrm{e}^{-rt}\phi_{0,1}(h),$$

其中 $\phi_{0,1}(x)$ 是标准正态分布函数. 对于第二项, 可能需要花一些精力,

$$\begin{aligned}\frac{S_0\mathrm{e}^{-rt}}{\sqrt{2\pi}}\int_{-\infty}^{h}\mathrm{e}^{by+a}\mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y&=\frac{S_0\mathrm{e}^{-rt}}{\sqrt{2\pi}}\int_{-\infty}^{h}\mathrm{e}^{-\frac{1}{2}(y^2-2by-2a)}\mathrm{d}y\\&=\frac{S_0\mathrm{e}^{-rt}}{\sqrt{2\pi}}\int_{-\infty}^{h}\mathrm{e}^{-\frac{1}{2}[(y-b)^2-b^2-2a]}\mathrm{d}y\\&=\frac{S_0\mathrm{e}^{-rt}}{\sqrt{2\pi}}\mathrm{e}^{\frac{1}{2}[b^2+2a]}\int_{-\infty}^{h}\mathrm{e}^{-\frac{1}{2}(y-b)^2}\mathrm{d}y\\&=S_0\mathrm{e}^{-rt+\frac{1}{2}b^2+a}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{h-b}\mathrm{e}^{-\frac{1}{2}y^2}\mathrm{d}y\\&=S_0\mathrm{e}^{-rt+\frac{1}{2}b^2+a}\phi_{0,1}(h-b).\end{aligned}$$

因此

$$P_{\infty}=K\mathrm{e}^{-rt}\phi_{0,1}(h)-S_0\mathrm{e}^{-rt+\frac{1}{2}b^2+a}\phi_{0,1}(h-b).$$

现在, 关于第二部分的指数我们有个令人高兴而吃惊的结果

$$\begin{aligned}-rt+\frac{1}{2}b^2+a&=-rt+\frac{1}{2}t\sigma_\nu^2+\frac{t}{2}(2r-\sigma_\nu^2)\\&=-rt+\frac{1}{2}t\sigma_\nu^2+rt-\frac{1}{2}t\sigma_\nu^2\\&=0.\end{aligned}$$

从而得到最后的公式为

$$P_{\infty}=K\mathrm{e}^{-rt}\phi_{0,1}(h)-S_0\phi_{0,1}(h-\sqrt{t}\sigma_\nu),$$

其中

$$h=\frac{1}{b}\left[\log\left(\frac{K}{S_0}\right)-a\right]=\frac{1}{\sqrt{t}\sigma_\nu}\left[\log\left(\frac{K}{S_0}\right)-tr+\frac{1}{2}t\sigma_\nu^2\right].$$

这就是关于欧式卖权的著名布莱克 – 舒尔斯公式.

我们可用卖权买权平价公式来得到相应的欧式买权价格. 回忆一下, 该公式告诉我们买权的价格为

$$C=P+S_0-K\mathrm{e}^{-rt}.$$

取 $T\to\infty$, 得到

$$\begin{aligned}C_{\infty}&=P_{\infty}+S_0-K\mathrm{e}^{-rt}\\&=K\mathrm{e}^{-rt}(\phi_{0,1}(h)-1)-S_0(\phi_{0,1}(h-\sqrt{t}\sigma_\nu)-1).\end{aligned}$$

因为

$$\phi_{0,1}(-t)=1-\phi_{0,1}(t),$$

所以买权的价格为

$$C_{\infty}=-K\mathrm{e}^{-rt}\phi_{0,1}(-h)+S_0\phi_{0,1}(\sqrt{t}\sigma_\nu-h).$$

令 $d_2=-h, d_1=d_2+\sigma\sqrt{t}$, 我们得到下面定理给出的公式.

定理 8(布莱克 – 舒尔斯期权定价公式)　对执行价为 K, 到期日为 t 的欧式期权来说, 有

$$\begin{aligned}C_{\infty}&=S_0\phi_{0,1}(d_1)-K\mathrm{e}^{-rt}\phi_{0,1}(d_2),\\P_{\infty}&=K\mathrm{e}^{-rt}\phi_{0,1}(-d_2)-S_0\phi_{0,1}(-d_1),\end{aligned}$$

其中 S_0 是标的股票的初始价格, σ 是瞬时波动率, $\phi_{0,1}$ 是标准正态分布函数, 且

$$d_1 = \frac{1}{\sigma\sqrt{t}}\left[\log\left(\frac{S_0}{K}\right) + t\left(r + \frac{1}{2}\sigma^2\right)\right],$$
$$d_2 = \frac{1}{\sigma\sqrt{t}}\left[\log\left(\frac{S_0}{K}\right) + t\left(r - \frac{1}{2}\sigma^2\right)\right] = d_1 - \sigma\sqrt{t},$$

这里 r 是无风险利率.

注意到该公式并不含有瞬时漂移系数. 事实上, 唯一“不知道”的量就是瞬时波动率 σ_ν 和无风险利率 r.

例 2 考虑一个欧式买权, 其标的股票的当前卖价为 \$100. 该期权还有 1 年就到期, 执行价为 \$100. 假设无风险利率为 0.05, 年波动率为 $\sigma = 0.15$. 试计算买权的价值.

解 这只是一个将数据代入公式的简单计算问题. 首先, 我们有

$$d_2 = \frac{1}{0.15}\left[0 + 0.05 - \frac{1}{2}(0.15)^2\right] \approx 0.2583$$

和

$$d_1 = d_2 + \sigma\sqrt{t} \approx 0.2583 + 0.15 = 0.4083.$$

因此, 利用计算器或者其他工具的帮助可计算 $\phi_{0,1}$ 的值, 得到

$$C = 100\phi_{0,1}(0.4083) - 100\mathrm{e}^{-0.05}\phi_{0,1}(0.5283) \approx \$8.596.$$

9.9 在实际中如何使用布莱克–舒尔斯公式：波动率微笑和波动率平面

至少可以这样说, 波动率为常数这个假设很不现实. 这就引出了有关布莱克 – 舒尔斯期权定价公式的优良性问题. 根据参数方面值得怀疑的假设, 已经做了大量的研究来确定如何使用布莱克 – 舒尔斯公式. 我们的目的就是简明地去讨论一种在实践中使用的方法.

波动率微笑 (volatility smiles)

我们已经说过波动率可由历史数据来估计. 但是, 在实践中, 使用布莱克 – 舒尔斯模型不是简单地代入波动率的一个估计值, 然后机械化地求出期权价格. 交易者经常用隐含波动率来代替. 粗略地说, 在布莱克 – 舒尔斯期权定价公式中所使用的波动率必须使公式在一个给定的时间内能反映出实际的市场价格.

定义 4　*考虑有特定市场价格 M 的欧式期权. 该期权的隐含波动率就是布莱克 – 舒尔斯期权定价公式所要求的波动率, 以使得该公式给出的价格刚好是 M.*

隐含波动率是有效的, 是市场对一只股票的布莱克 – 舒尔斯波动率的看法. 隐含波动率可根据布莱克 – 舒尔斯公式通过数值方法计算出 (即反复猜测).

就像刚才看到的, 且还有更深入的证据表明布莱克 – 舒尔斯公式是不完美的, 如果去计算只是执行价不同的期权的隐含波动率, 可得到不同的值. 图 5 给出了隐含波动率和执行价之间的典型图形, 这就是所谓的波动率微笑.

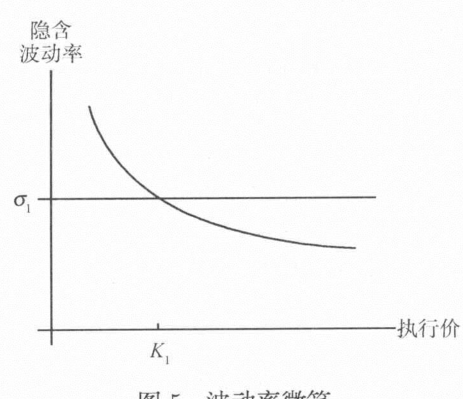


图 5　波动率微笑

现在, 假设一只特定股票的历史波动率的估计值为 σ_1, 则布莱克 – 舒尔斯公式给出一个期权价格, 该价格与执行价为 K_1 的期权的市场价格相匹配.

对于更高的执行价, 这对应着虚值买权和实值卖权, 市场价格对应着一个比 σ_1 小的波动率. 现在, 布莱克 – 舒尔斯公式给出的价格直接随波动率变化, 因而较小的波动率对应着较低的价格. 因此, 市场价格低于布莱克 – 舒尔斯历史价格. 换种方式说, 在使用布莱克 – 舒尔斯公式时, 如果使用的波动率为基于历史数据的一个常数波动率, 那么实值买权和虚值卖权的价格相对市价来说就较低。

波动率平面 (volatility surfaces)

为了理解如何在实际中使用布莱克 – 舒尔斯公式, 我们来考虑表 1, 该表给出了一个波动率平面的数据, 即不同到期日和执行价所对应的隐含波动率值. 各列表示不同的执行价, 用股价的百分数表示 (该表中的数据仅供举例说明).

现在, 一个交易员想对一只期权定价, 该期权的到期日和执行价不在表 1 中, 他可用插值法来得到隐含波动率. 例如, 考虑一个还有 9 个月就到期, 执行价为股价的 95% 的期权. 在 6 个月和 1 年的隐含波动率之间用线性插值得到波动率为 $(14.3 + 15)/2 = 14.65$. 可用该波动率代入到布莱克 – 舒尔斯公式中而计算出所考

虑的期权价格.

表 1 波动率平面数据

	90%	95%	100%	105%	110%
1 个月	14	13	12	11.3	14.4
3 个月	14.2	14.1	13.6	13.8	14.1
6 个月	14.5	14.3	14.2	14.5	14.6
1 年	15.1	15	14.6	14.7	14.8
2 年	16.2	16.1	16	16.1	16.2

9.10 红利如何影响布莱克 – 舒尔斯公式的使用

布莱克 – 舒尔斯期权定价公式假定标的股票不分红. 我们在该节中简要讨论如何处理红利.

首先了解一点有关红利的背景知识. 有四个重要日期与红利相关: 宣告日 (declaration date), 就是董事会宣布发放红利的日期; 登记日 (record date), 就是登记员登记当前持股人的名单以确定红利分发给谁的日期. 非常重要的一点是投资者必须在登记日登记, 否则他不会收到红利; 支付日 (payment date), 就是红利支付的日期.

现在, 购买普通股需花 3 天时间来清算. 这被称为规范的结算方式. 因此, 拥有者必须在 3 天前 (相对登记日来说) 就已经购买了股票, 才能够在登记日被登记下来, 因此才有资格获取红利. 称该天后的第一天为除权日 (ex-dividend date). 例如, 如果在某一星期五宣布给登记的持有者派发红利, 那么纽约股票交易所 (它规定 NYSE 中股票的除权日) 会宣布下周三开盘时的股价为除权后的价格.

我们注意到, 当股票除权后, 典型的情形是股价会下跌, 下跌的数量与分红的数量相似. 从提前知道红利从而红利嵌在股价里这个观点来看是合理的.

现在, 可以认为标的股票派发红利的欧式期权由两个分离的过程组成: 一个表示股价本身的风险过程, 另一个表示现金红利支付的无风险过程. 因此, 为了定价这样的期权, 首先我们需将未来的所有红利支付折现到现在. 如果该值为 d, 那么我们认为组成为初始价值为 $S_0 - d$ 的有风险股票和初始价值为 d 的无风险资产. 然后就可以对有风险股票运用布莱克 – 舒尔斯公式.

练 习 9

1. 一只股票的初始价格为 \$50. 在周期为 5 天的一个周期内, 股票每天最后的价格为 \$49.82, \$50.02, \$49.69, \$49.34, \$50.10. 试估计瞬时漂移系数和波动率.

2. 考虑一只以股票为标的的欧式买权, 其当前售价为 \$80. 执行价为 \$80 的期

权还有 1 年就要到期. 假设无风险利率为 0.05, 年波动率为 $\sigma = 0.1$. 计算该买权的价值.

3. 考虑一只以股票为标的的欧式卖权, 其当前售价为 \$80. 执行价为 \$80 的期权还有 1 年就要到期. 假设无风险利率为 0.04, 年波动率为 $\sigma = 0.15$. 计算该卖权的价值.

4. 证明：

$$\mathrm{var}_p(X_{p,T,i}) = 1.$$

5. 证明：

$$\mathcal{E}_p(Q_s) = \sigma_s \frac{t}{\sqrt{\Delta t}} \frac{p-s}{\sqrt{s(1-s)}},$$

$$\mathrm{var}_p(Q_s) = \sigma_s^2 t \frac{p(1-p)}{s(1-s)}.$$

6. 设随机变量 X 服从对数正态分布, 即 $Y = \log X$ 是均值为 a 方差为 b^2 的正态随机变量. 证明：

a) $\mathcal{E}(X) = \mathrm{e}^{a+\frac{1}{2}b^2}$.

b) $\mathrm{var}(X) = \mathrm{e}^{2a+b^2}(\mathrm{e}^{b^2} - 1)$.

对随机变量 $X = \dfrac{S_t}{S_0}$ 应用上面结果, 请推出

c) $\mathcal{E}_\pi(S_t) = S_0 \mathrm{e}^{rt}$.

d) $\mathrm{var}_\pi(S_t) = (S_0 \mathrm{e}^{rt})^2 (\mathrm{e}^{\sigma_\nu^2 t} - 1)$.

7. 证明：函数 $f(x) = \max\{K - S_0 \mathrm{e}^x, 0\}$ 在 $\mathbb{R}$ 上连续且有界.

8. 证明：对标准正态分布函数 $\phi_{0,1}$, 有 $\phi_{0,1}(-t) = 1 - \phi_{0,1}(t)$. 提示：利用标准正态分布的密度函数的曲线图画个图像.

9. 用洛必达法则证明：

$$\lim_{T\to\infty} \pi_T = \lim_{\Delta t \to 0} \left(\frac{\mathrm{e}^{(r-\mu)\Delta t + \frac{p}{\sqrt{pq}}\sigma\sqrt{\Delta t}} - 1}{\mathrm{e}^{\frac{1}{\sqrt{pq}}\sigma\sqrt{\Delta t}} - 1} \right) = p.$$

提示：用 $x = (\Delta t)^2$.

10. 假设在 $T \to \infty$ 的情形中存在一个鞅测度, 并且资产定价基本定理在该情形下成立. 如果我们记鞅测度为 Π, 那么

$$S_0 = \mathrm{e}^{-rt} \mathcal{E}_\Pi(S_t) = \mathrm{e}^{-rt} \mathcal{E}_\Pi(S_0 \mathrm{e}^{\mu_\nu t + \sigma_\nu \sqrt{t} Z_t}).$$

通过计算最后一个式子来证明

$$\mu_\nu = r - \frac{1}{2}\sigma_\nu.$$

11. 设随机变量 N_U 代表股价在 CRR 模型的整个周期中上涨的次数. 证明：

$$S_{t,T} = S_0 \mathrm{e}^{N_U(\log u - \log d) + T \log d}.$$

所以

$$H_{t,T} = N_U(\log u - \log d) + T \log d.$$

证明：

$$\mathcal{E}(H_{t,T}) = T\nu(\log u - \log d) + T \log d,$$
$$\mathrm{var}(H_{t,T}) = T\nu(1-\nu)(\log u - \log d)^2.$$

12. 证明标准化的随机变量 $H_{t,T}$ 有

$$H^*_{t,T} = N^*_U.$$

因此, 随机变量 $H^*_{t,T}$ 是标准的二项式分布随机变量.

13. 利用布莱克－舒尔斯公式证明卖权和买权的价值随着波动率 σ 的增大而增大. 请看多头卖权和多头买权的利润曲线图, 解释为什么这是合理的. 对于股票持有者会有相同的效果吗?

14. 请证明执行价为 K 到期日为 t 的欧式买权到期时处于平价或实值的概率为

$$\phi_{0,1}\left(\frac{\log(S_T/K) - t(r - \sigma_\nu^2/2)}{\sigma\sqrt{t}}\right),$$

其中 $\phi_{0,1}$ 是标准正态分布函数.

第 10 章　最优停时和美式期权

至今为止, 我们建立的模型 (包括布莱克 – 舒尔斯期权定价模型) 都是为了定价欧式期权, 一种只能在到期日才能执行的期权. 然而, 在现实世界中, 很多股票期权是美式的. 在这一章, 我们想来看看美式期权的定价问题.

美式期权比欧式期权复杂很多, 这是因为它没有丢掉什么但增加了一个特征 —— 能在购买日与到期日之间的任何时刻执行美式期权. 这是一个重要的改变, 因为没有办法知道它会在未来什么时候被执行. 一个投资者不会对他的经纪人说: “如果股价会下跌到 $50 以下, 那么在下跌之前把股票卖掉” .

美式期权模型中涉及的数学知识与以前的显著不同, 比我们到目前为止所碰到的都要复杂些.

10.1　一个例子

为了帮助讨论, 我们从一个简单例子入手, 在后面还会用到该例子.

例 1　图 1 是一个三期 CRR 模型中的股价 (和期权损益) 变化图.

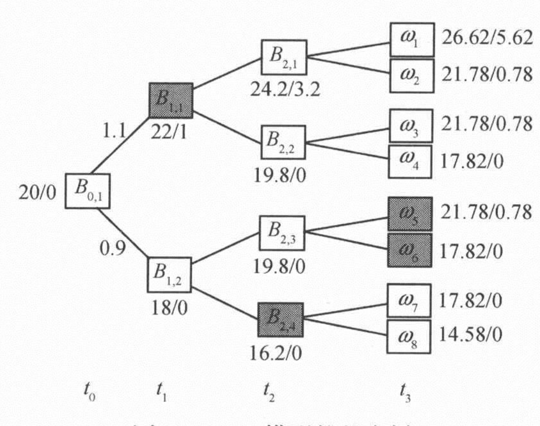


图 1　CRR 模型的状态树

我们假设 $r = 0$. 注意到在该模型中

$$T = 3, \quad u = 1.1, \quad d = 0.9,$$

且鞅测度为

$$\pi = \frac{1-d}{1-u} = \frac{0.1}{0.2} = \frac{1}{2}.$$

最后, 令 C 是执行价为 $K=21$ 的美式买权 C.

在开始之前先提一下有关记号的意义. 对任意随机变量 X, 很方便定义容易记的符号 $[X \in A]$ 为

$$[X \in A] = \{\omega \in \Omega | X(\omega) \in A\}.$$

10.2 模　型

一般地, 我们考虑的是无套利的离散时间模型, 因此存在一个等价鞅测度 Π, 例 1 就是这种情形. 考虑投资美式期权 (也称美式要求权). 持有者可在模型中的任何时刻 $t_0 < t_1 < \cdots < t_T$ 执行期权.

10.3 损　益

期权在任何时刻 t_k 的损益是一个随机变量 Y_k. 可以假定 (Y_k) 关于流 $\mathbb{F} = (\mathcal{P}_k)$ 是适应的. 在我们的例子中, 美式买权 C 的损益为

$$Y_3 = \max\{S_3 - 21, 0\} = \begin{cases} 5.62, & \omega = \omega_1, \\ 0.78, & \omega = \omega_2, \omega_3, \omega_5, \\ 0, & \text{其他}, \end{cases}$$

$$Y_2 = \max\{S_2 - 21, 0\} = \begin{cases} 3.2, & \omega \in B_{2,1}, \\ 0, & \text{其他}, \end{cases}$$

$$Y_1 = \max\{S_1 - 21, 0\} = \begin{cases} 1, & \omega \in B_{1,1}, \\ 0, & \text{其他} \end{cases}$$

和

$$Y_0 = 0.$$

这些损益如图 1 所示.

10.4 停　时

可以将什么时候该执行期权的决策看成一个有特殊性质的随机变量, 叫做停时. 想法很简单：一个停时就是一个规则, 也就是说, 它是这样的一个随机变量, 指出 t_k 时 Ω 中的哪些结果要求我们在 t_k 时应该执行期权. 我们把这些结果的集合看作 t_k

时的停止事件 (stopping event). 但是对最终时刻 t_T 来说, 停止事件由在 t_T 时要么执行要么放弃的结果组成.

很明显, 对一个停时的唯一要求是：在 t_k 时我们必须能够说出哪些结果属于 t_k 时的停止事件. 这是非常重要的一点. 例如, 基于 t_3 时所发生的我们不能说 t_2 时的执行事件. 这与叫经纪人在股价下跌到 50 以下之前把股票卖掉相类似.

定义 1 (有界) 停时是一个随机变量

$$\tau : \Omega \to \{0, \cdots, T\},$$

它的取值为 0 到 T 的整数. 而且, 要求 t_k 时的停止事件

$$[\tau = k] = \{\omega \in \Omega | \tau(\omega) = k\}$$

属于由 $\mathcal{P}_k$ 生成的代数 $\mathcal{A}(\mathcal{P}_k)$, 对所有的 $k = 0, \cdots, T$ 成立. 用 $\mathcal{S}_{k,T}$ 表示取值范围为 $\{k, \cdots, T\}$ 的所有停时. 这些停时在 t_k 之前不会停止.

我们考虑停时的一个非常重要的例子.

例 2 例如, 投资者一般会对他的经纪人说:“当股价达到 \$50 或者更高就卖掉股票”. 该规则就是一个停时. 事实上, 从数学上看就是股票价格过程 (S_k) 首次进入集合 $[50, \infty)$ 的时间 (first entry time). 正式定义如下：

$$\tau(\omega) = \begin{cases} \min\{k | S_k \geqslant 50\}, & \text{如果}\{k | S_k \geqslant 50\} \neq \varnothing, \\ T, & \text{其他}. \end{cases}$$

不难证明 τ 确实是一个停时. 因为如果 $k < T$, 那么有

$$\begin{aligned} [\tau = k] &= \{\omega | S_k \geqslant 50, \text{ 但是对 } j < k \text{ 有 } S_j < 50\} \\ &= [S_0 < 50] \cap \cdots \cap [S_{k-1} < 50] \cap [S_k \geqslant 50]. \end{aligned}$$

但是价格 S_i 是 $\mathcal{P}_i$ 可测的且 $(\mathcal{P}_i)$ 是一个流, 所以可推出每个事件 $[S_i < 50]$ 和事件 $[S_k \geqslant 50]$ 都在最大的代数 $\mathcal{A}(\mathcal{P}_k)$ 里. 这就是对停时的要求条件. 最后, 对 $k = T$ 有

$$[\tau = T] = [S_0 < 50] \cap \cdots \cap [S_{T-1} < 50] \in \mathcal{A}(\mathcal{P}_T).$$

因此, τ 是一个停时. 注意对其他数值而不是 50 可进行同样的讨论.

事实上, 也可以证明任何 Borel 集 B 的首次进入时是停时. 例如, 集合

$$B = (-\infty, 17) \cup (20, \infty)$$

对应着股价首次下跌到 17 以下或者上涨到 20 以上. 图 1 的阴影部分指出了 B 的首次进入时的停止事件.

10.5 损益的停止过程

这里有一种情景. 假想有一位投资者在 t_0 时拥有一份美式期权. 该投资者坐观集合 $S_{0,T}$ 里的所有可能的停时 (这至少在理论上是可行, 因为这样的停时只有有限个).

假设投资者已经决定采用一个特殊的停时 $\tau \in \mathcal{S}_{0,T}$ 来确定什么时候执行期权. 我们稍后将讨论该决策是如何做出的. 事实上, 这也是该章的主要内容.

想象投资者在 t_0 时给他的经纪人打电话并告诉他的停止规则 τ, 这是有帮助的. 从这一刻出发, 经纪人就可以一直进行下去而不需要打扰投资者. 特别地, 在每个时刻 t_k 经纪人核查当前的经济状态是否在停止事件 $[\tau = k]$ 内. 他能这样做的原因是因为 $[\tau = k] \in \mathcal{A}(\mathcal{P}_k)$ 刚好是 $\mathcal{P}_k$ 中块的并, 因此经纪人知道哪个时候哪个块代表当前的状态. 如果当前状态在 $[\tau = k]$ 内, 那么经纪人就执行期权, 否则就不执行.

10.6 美式期权的停止价值

为了确定如何去选择一个停时, 首先我们必须讨论有关停时的任意选择的序列. 假设投资者决定采用一个特殊停时 $\tau \in \mathcal{S}_{0,T}$, 那么对任意的 $\omega \in \Omega$, 在时刻 $t_{\tau(\omega)}$ 执行期权, 这时的损益为

$$Y_{\tau(\omega)}(\omega),$$

习惯用 Y_τ 来表示该函数. 因此

$$[Y_\tau](\omega) = Y_{\tau(\omega)}(\omega).$$

称随机变量 Y_τ 为过程 (Y_k) 在停时 τ 下的最终价值.

例 3 接例 1, 考虑图 1 所示的停时 τ. 它是首次进入 B 的时间, 其中

$$B = (-\infty, 17) \cup (20, \infty).$$

(折现的) 最终价值 Y_τ 为

$$Y_\tau(\omega) = \begin{cases} 1, & \text{如果}\omega \in \{\omega_1, \omega_2, \omega_3, \omega_4\}, \\ 0.78, & \text{如果}\omega = \omega_5, \\ 0, & \text{如果}\omega \in \{\omega_6, \omega_7, \omega_8\}. \end{cases}$$

有关折现的细节

我们现在讨论折现的细节问题. 如果 X_k 是任意一个过程, $\tau \in \mathcal{S}_{0,T}$ 是一个停时, 那么最终价值就是 X_τ. 为了得到该价值的折现值, 我们必须将每个 $(X_\tau)(\omega) = X_{\tau(\omega)}(\omega)$ 按一个合适的量进行折现

$$\overline{X_{\tau(\omega)}(\omega)} = \frac{X_{\tau(\omega)}(\omega)}{B_{\tau(\omega)}(\omega)} = \left(\frac{X_\tau}{B_\tau}\right)(\omega).$$

因此, 令

$$\overline{X_\tau} = \frac{X_\tau}{B_\tau}.$$

注意, 在例 1 中假设无风险利率为 0, 因此没有相应的折现问题.

10.7　美式期权的初始价值或在时刻 t_0 该做什么

在时刻 t_0, 停时的可能选择为集合 $\mathcal{S}_{0,T}$ 中的元素. 如果投资者采用一个特殊的停时 $\tau \in \mathcal{S}_{0,T}$ 来停止损益过程, 那么他得到的最终价值为 Y_τ. 但是, 这里还有一点点微妙关系, 也就是对每个 $\omega \in \Omega$, 损益 $Y_{\tau(\omega)}(\omega)$ 在 $t_{\tau(\omega)}$ 时发生. 为了保持自融资交易策略成立, 我们假定投资者不从他的经纪账户中移走资金, 而一直坚持到模型的最后时刻 t_T, 从而允许 $t_{\tau(\omega)}$ 时的损益从时期 $\tau(\omega)$ 到 T 以无风险利率增长. 这就导致最终的损益为

$$\mathrm{e}^{r(T-\tau(\omega))}Y_{\tau(\omega)}(\omega) = \frac{B_T}{B_{\tau(\omega)}}Y_{\tau(\omega)}(\omega),$$

其中 $B_k = \mathrm{e}^{rk}$ 是折现因子.

因此, 最后时刻 t_T 的损益真正来自停时 τ 的部分为

$$\frac{B_T}{B_\tau}Y_\tau = B_T\overline{Y_\tau},$$

其中

$$\overline{Y_\tau} = \frac{Y_\tau}{B_\tau}.$$

另一方面, 每个停时将美式期权转变为一列有保证的损益序列, 在 t_k 时的损益为

$$Y_k 1_{[\tau=k]}.$$

损益流的最终价值为

$$\sum_{k=0}^{T}\frac{B_T}{B_k}Y_k 1_{[\tau=k]} = \frac{B_T}{B_\tau}Y_\tau = B_T\overline{Y_\tau}.$$

美式期权的初始价值

现在, 我们来考虑投资者应该做些什么来确定在 t_0 时采用哪个停时. 记住一点, 投资者可能在 t_1 时改变他的想法, 但是我们不必担心这类问题.

如前面提到的那样, 为了确定 t_0 时的最优停时, 投资者可以观察有限集 $\mathcal{S}_{0,T}$ 里的所有可能的停时. 首先, 看起来较合理的是投资者应该选择使最终损益最大化的停时, 即

$$\max_{\tau\in\mathcal{S}_{0,T}}\{B_T\overline{Y_\tau}\}.$$

但是, 支付 $B_T\overline{Y_\tau}$ 是函数 (随机变量), 故不能保证存在一个停时对所有的 $\omega\in\Omega$ 来说都是最优的. 确实, 这是很不可能的.

因此换一种方法是有必要的: 我们可以最大化损益的初始价值. 假设最终损益 $B_T\overline{Y_\tau}$ 是可达到的, 则存在自融资交易策略 Φ, 使得

$$\mathcal{V}_T(\Phi)=B_T\overline{Y_\tau},$$

且最终损益的无套利价格为 (根据鞅测度条件)

$$\begin{aligned}V_0(\tau)&=\mathcal{I}(B_T\overline{Y_\tau})\\&=\mathcal{V}_0(\Phi)\\&=\frac{1}{B_T}\mathcal{E}_\Pi(\mathcal{V}_T(\Phi))\\&=\frac{1}{B_T}\mathcal{E}_\Pi(B_T\overline{Y_\tau})\\&=\mathcal{E}_\Pi(\overline{Y_\tau}),\end{aligned}$$

这里 $V_0(\tau)$ 就是美式期权在 τ 下的初始价值. 我们已经指出, 假定期权的拥有者遵循停时 τ, 则未定权益的无套利价格一定是 $V_0(\tau)$.

现在, 投资者在时刻 t_0 可以选择一个使期权的初始价值最大化的停时, 这是因为不同停时下的初始价值都是常数. 因此定义

$$V_0=\max_{\tau\in\,\mathcal{S}_{0,T}}V_0(\tau)=\max_{\tau\in\,\mathcal{S}_{0,T}}\mathcal{E}_\Pi(\overline{Y_\tau}).$$

从而可以假设投资者会选择一个停时 τ^*, 该停时满足

$$\mathcal{E}_\Pi(\overline{Y_{\tau^*}})=V_0.$$

这就是使期权初始价值最大化的停时, 或者等价地, 使鞅测度下的最终期望损益 $\mathcal{E}_\Pi(\overline{Y_\tau})$ 最大化. 称这样的停时为最优停时. 由于不需要考虑未来, 这可能是能做到的最好办法了.

让我们站在这个位置上来更近地看看套利. 如以前看到的, 对于一个有可达到的最终损益 $\mathcal{V}_T(\Phi)$ 的欧式期权来说, 它的无套利初始价值一定是 $\mathcal{V}_0(\Phi)$. 如果不是, 那么投资者就可以从 Φ 和期权二者中购买较便宜的而卖出较贵的, 而立即有一个正的收入. 到期末 t_T 时, 欧式期权的损益为 $\mathcal{V}_T(\Phi)$, 因此期末时两个头寸 (一个多头, 一个空头) 刚好对冲掉, 而投资者可以获得初始利润的未来价值.

但是, 对于具有下面损益的美式期权来说

$$\mathcal{V}_T(\Phi) = B_T\overline{Y_{\tau^*}},$$

情况就并不那么简单了, 这是因为美式期权的出售者并不知道最终损益是多少 (购买者也不知道, 但是他至少对损益价值有一些控制).

我们可以说, 如果期权的价格 A 小于 $\mathcal{V}_0(\Phi)$, 那么对投资者来说就存在套利, 他可以购买较便宜的期权而做空较贵的交易策略 Φ. 与欧式期权一样, 拥有者将期权持有到期, 这就可以获得初始利润而在期末时拥有相互抵消的头寸.

然而, 如果期权的价格 A 大于 $\mathcal{V}_0(\Phi)$, 那么通过卖出期权 (同时买进 Φ) 可以立即获得利润. 但是期权出售者不能控制期权, 不能确定期权拥有者不会实现一个比 Φ 更高的支付. 这里关键的一点就是, 对欧式期权起作用的复制策略 Φ 对美式期权并不起作用.

因此, 由无套利讨论可得出 $A \geqslant \mathcal{V}_0(\Phi)$. 如果在 t_0 时不能保证存在一个最优停时, 在该停时下得到的损益大于 $\mathcal{V}_T(\Phi)$, 那么谁还会愿意支付大于 $\mathcal{V}_0(\Phi)$ 来购买美式期权? 故我们得到一个结论, 美式期权的公平价格 (或多或少) 是 $\mathcal{V}_0(\Phi)$.

例 4　再次考虑例 1, 我们已经知道首次进入

$$B = (-\infty, 17) \cup (20, \infty)$$

的最终损益为

$$Y_\tau(\omega) = \begin{cases} 1, & \text{如果}\omega \in \{\omega_1, \omega_2, \omega_3, \omega_4\}, \\ 0.78, & \text{如果}\omega = \omega_5, \\ 0, & \text{如果}\omega \in \{\omega_6, \omega_7, \omega_8\}. \end{cases}$$

因此

$$\mathcal{E}_\Pi(\overline{Y_\tau}) = \frac{1}{2}\cdot 1 + \frac{1}{8}\cdot 0.78 + \frac{3}{8}\cdot 0 = 0.5975.$$

考虑如下定义的停时 σ,

$$\sigma(\omega) = \begin{cases} 2, & \text{如果}\omega \in \{\omega_1, \omega_2, \omega_7, \omega_8\}, \\ 3, & \text{其他}. \end{cases}$$

我们将 σ 是一个停时的证明留给读者. 在这种情形下, (折现的) 最终价值为

$$Y_\sigma(\omega)=\begin{cases}3.2, & 如果\omega\in\{\omega_1,\omega_2\},\\ 0.78, & 如果\omega=\omega_3,\omega_5,\\ 0, & 如果\omega\in\{\omega_4,\omega_6,\omega_7,\omega_8\}.\end{cases}$$

因此

$$\mathcal{E}_\Pi(\overline{Y_\sigma})=\frac{1}{4}\cdot 3.2+\frac{1}{4}\cdot 0.78=0.995.$$

所以 σ 是比 τ 更好的停时. 事实上, 我们将看到 σ 是一个最优停时.

10.8 t_k 时该做什么

现在假设投资者在 t_k 时还没有执行美式期权. 只需做一些必要的修正, 前面的讨论还是有效的. 特别地, 目前的停时必须从 $\mathcal{S}_{k,T}$ 中选出, 这是因为投资者此时不可能在 t_k 以前的任何时刻执行期权.

如果投资者采用一个特殊的停时 $\tau\in\mathcal{S}_{k,T}$ 来停止损益过程, 那么他会实现价值 Y_τ, 其在最后时刻 t_T 的价值为

$$\frac{B_T}{B_\tau}Y_\tau=B_T\overline{Y_\tau}.$$

假设最终损益 $B_T\overline{Y_\tau}$ 是可达到的, 则存在自融资交易策略 Φ 满足

$$\mathcal{V}_T(\Phi)=B_T\overline{Y_\tau},$$

且最终损益在 t_k 时的无套利价格为

$$\begin{aligned}V_k(\tau)&=\mathcal{I}_k(B_T\overline{Y_\tau})\\&=\mathcal{V}_k(\Phi)\\&=\frac{B_k}{B_T}\mathcal{E}_\Pi(\mathcal{V}_T(\Phi)|\mathcal{F}_k)\\&=\mathcal{E}_\Pi(B_k\overline{Y_\tau}|\mathcal{F}_k),\end{aligned}$$

这里 $V_k(\tau)$ 就是美式期权 t_k 时在 τ 下的价值.

现在, 我们又假定投资者在时刻 t_k 做出可能的最优停止决策, 在该情形下等于去选择一个停时 τ^* 使得 $V_k(\tau^*)$ 是最大值. 相应地, 定义 V_k 如下

$$V_k=\max_{\tau\in\,\mathcal{S}_{k,T}}V_k(\tau)=\max_{\tau\in\,\mathcal{S}_{k,T}}\mathcal{E}_\Pi(B_k\overline{Y_\tau}|\mathcal{F}_k).$$

并称停时 τ^* 是最优的, 如果

$$\mathcal{E}_\Pi(B_k\overline{Y_{\tau^*}}|\mathcal{F}_k) = V_k.$$

我们称 (V_k) 为美式期权的价值过程.

我们花一点时间来比较 t_0 时的决策过程, 即相应地最大化

$$V_0 = \max_{\tau\in \mathcal{S}_{0,T}} V_k(\tau) = \max_{\tau\in \mathcal{S}_{0,T}} \mathcal{E}_\Pi(\overline{Y_\tau})$$

和 t_k 时的决策过程, 即相应地最大化

$$V_k = \max_{\tau\in \mathcal{S}_{k,T}} V_k(\tau) = \max_{\tau\in \mathcal{S}_{k,T}} \mathcal{E}_\Pi(B_k\overline{Y_\tau}|\mathcal{F}_k).$$

在后面这种情形里, 我们在一个较小的集合上取最大值, 因为

$$\mathcal{S}_{k,T} \subseteq \mathcal{S}_{0,T}.$$

因此这倾向于使最大值相对较小. 另一方面, t_k 时我们是在具有更多信息的基础上进行最大化的, 这就是说, 我们最大化条件期望 $\mathcal{E}_\Pi(B_k\overline{Y_\tau}|\mathcal{F}_k)$, 故这又倾向于使最大值相对较大. 因此, 我们遇到了两股相抵触的力量, 对于哪个更大我们一无所知.

定义 2 损益过程为 $(Y_k|\ k=0,\cdots,T)$ 的美式期权的无套利价格过程是

$$V_k = \max_{\tau\in \mathcal{S}_{k,T}} \mathcal{E}_\Pi(B_k\overline{Y_\tau}|\mathcal{F}_k).$$

折现价值过程为

$$\overline{V_k} = \max_{\tau\in \mathcal{S}_{k,T}} \mathcal{E}_\Pi(\overline{Y_\tau}|\mathcal{F}_k).$$

称折现价值过程 $(\overline{V_k})$ 为折现损益过程 $(\overline{Y_k})$ 的 Snell 包络.

定义 3 停时 τ^* 在 $[k,T]$ 上是最优的, 如果它最大化了期望折现损益过程 $(\overline{Y_k})$, 即

$$\mathcal{E}_\Pi(\overline{Y_{\tau^*}}|\mathcal{F}_k) = \overline{V_k} = \max_{\tau\in \mathcal{S}_{k,T}} \mathcal{E}_\Pi(\overline{Y_\tau}|\mathcal{F}_k).$$

也就是说, 如果 τ^* 实现了最佳期望折现损益.

10.9 最优停时和 Snell 包络

为了使符号简单化, 我们利用任意的非折现过程来学习 Snell 包络. 唯一的区别就是我们是否需要包括上划线.

定义 4 如果 $\mathbb{Z} = (Z_k | k = 0, \cdots, T)$ 是一个随机过程, 那么称如下定义的过程 $\mathbb{U} = (U_k)$ 为 $\mathbb{Z}$ 的 Snell 包络

$$U_k = \max_{\tau \in \mathcal{S}_{k,T}} \mathcal{E}_\Pi(Z_\tau | \mathcal{F}_k).$$

如果停时 τ^* 在 $[k, T]$ 上实现了最大值, 即

$$\mathcal{E}_\Pi(Z_{\tau^*} | \mathcal{F}_k) = U_k = \max_{\tau \in \mathcal{S}_{k,T}} \mathcal{E}_\Pi(Z_\tau | \mathcal{F}_k).$$

那么称 τ^* 为 $\mathbb{Z}$ 在 $[k, T]$ 上的最优停时.

因此, 如果 $\mathbb{Z}$ 是美式期权的折现损益过程, 那么 Snell 包络 $\mathbb{U}$ 就是期权的折现价值过程.

稍后, 当我们有一些附加工具可使计算更简单时, 我们将计算例 1 中的损益过程 (Y_k) 的 Snell 包络.

10.10 最优停时的存在性

关于最优停时的第一个重要问题就是它是否存在. 在 $k = 0$ 时, 最优停时显然存在, 这是因为我们只需简单地最大化一个常数集 $\mathcal{E}_\Pi(Z_\tau)$. 但是当 $k > 0$ 时, 我们需要最大化的是非常数的函数 $\mathcal{E}_\Pi(Z_\tau | \mathcal{F}_k)$.

定理 1 在任何区间 $[k, T]$ 内, $\mathbb{Z}$ 的最优停时存在.

证明 请回忆一下, 对

$$\mathcal{P}_k = \{B_{k,1}, \cdots, B_{k,c}\},$$

条件期望的定义如下

$$\mathcal{E}_\Pi(Z_\tau | \mathcal{P}_k) = \sum_{u=1}^{c} \mathcal{E}_\Pi(Z_\tau | B_{k,u}) 1_{B_{k,u}}.$$

因此, 对每个 $B_{k,u}$, 随机变量 $\mathcal{E}_\Pi(Z_\tau | \mathcal{P}_k)$ 在 $B_{k,u}$ 上等于常数 $\mathcal{E}_\Pi(Z_\tau | B_{k,u})$. 所以我们可以找到一个停时 $\tau_{k,u}$ 来最大化这些常数, 即

$$\mathcal{E}_\Pi(Z_{\tau_{k,u}} | B_{k,u}) = \max_{\tau \in \mathcal{S}_{k,T}} \mathcal{E}_\Pi(Z_\tau | B_{k,u}).$$

考虑随机变量

$$\tau_k^* = \sum_{u=1}^{c} \tau_{k,u} 1_{B_{k,u}}.$$

它使条件期望在 $\mathcal{P}_k$ 的每个块上达到最大. 为了说明 τ^* 是 $\mathcal{S}_{k,T}$ 里的一个停时, 注意到 $\tau^* \geqslant k$, 且对任意的 $h \geqslant k$, 有

$$[\tau_k^* = h] = \bigcup_{v=1}^{c}([\tau_k^* = h] \cap B_{k,v}) = \bigcup_{v=1}^{c}([\tau_{k,v} = h] \cap B_{k,v}) \in \mathcal{A}(\mathcal{P}_h),$$

这是停时的判别条件.

现在, 如果 $\omega \in B_{k,v}$, 那么

$$[Z_{\tau_k^*}](\omega) = Z_{\tau_k^*(\omega)}(\omega) = Z_{\tau_{k,v}(\omega)}(\omega).$$

从而

$$Z_{\tau_k^*} = \sum_{u=1}^{c} Z_{\tau_{k,u}} 1_{B_{k,u}}.$$

因此, 对任何 $\tau \in \mathcal{S}_{k,T}$, 有

$$\begin{aligned}
\mathcal{E}_\Pi(Z_{\tau_k^*}|\mathcal{P}_k) &= \sum_{u=1}^{c} \mathcal{E}_\Pi(Z_{\tau_{k,u}} 1_{B_{k,u}}|\mathcal{P}_k) \\
&\geqslant \sum_{u=1}^{c} \mathcal{E}_\Pi(Z_\tau 1_{B_{k,u}}|\mathcal{P}_k) \\
&= \sum_{u=1}^{c} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k) 1_{B_{k,u}} \\
&= \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k) \sum_{u=1}^{c} 1_{B_{k,u}} \\
&= \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k).
\end{aligned}$$

这就是判定最优停时所需要的证明.

我们也来证明 Snell 包络是 $\mathbb{F}$ 适应的.

定理 2　*Snell 包络 (U_k) 是 $\mathbb{F}$ 适应的.*

证明　随机变量 U_k 是有限个随机变量 $\mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k)$ 中的最大值, 由于每个 $\mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k)$ 是 $\mathcal{P}_k$ 可测的, 因此 U_k 也是的.

10.11　Snell 包络的刻画

再一次来考虑投资者的境况, 他在 t_k 时需要从 $\mathcal{S}_{k,T}$ 的所有停时中作出一个决策. 当计算一个最优停时时, 他可以基于当前时刻 t_k 的停止事件将候选项 $\mathcal{S}_{k,T}$ 分成三个子集.

投资者可能现在就能很容易地就作出停止的决定 (在 t_k 时), 然后就完成了任务, 也可以决定在 t_k 时不停止而不管环境如何, 或者他的决定依赖一个停时, 可能现在停止也可能稍后才停止, 这取决于经济的状态. 用符号表示, 集合 $\mathcal{S}_{k,T}$ 为不相交集合的并

$$\mathcal{S}_{k,T} = \mathcal{S}_{k+1,T} \cup \mathcal{S}_{k,k} \cup \mathcal{S}^*_{k,T},$$

对应如下:

1) 现在不停止, 即在 t_{k+1} 或更后停止

$$\mathcal{S}_{k+1,T} = \{\tau \in \mathcal{S}_{k,T} |\ [\tau = k] = \varnothing\}.$$

2) 现在停止 (在时刻 t_k)

$$\mathcal{S}_{k,k} = \{k1_\Omega\} = \{\tau \in \mathcal{S}_{k,T} |\ [\tau = k] = \Omega\}.$$

3) 可能在 t_k 停止也可能在更后停止

$$\mathcal{S}^*_{k,T} = \{\tau \in \mathcal{S}_{k,T} |\ [\tau = k] \neq \varnothing, \Omega\}.$$

我们想要指出, 不需要考虑类型 3) 的停时而可以计算出 Snell 包络. 注意, 我们不是说在 3) 这种情况下不存在最优停时, 而仅仅是说计算 U_k 时不需要考虑类型 3) 的停时.

从数学观点看上面的陈述得到 Snell 包络满足

$$U_k = \max\{Z_k, \max_{\tau \in \mathcal{S}_{k+1,T}} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k)\}.$$

注意, 现在仅需从集合 $\mathcal{S}_{k+1,T}$ 上来考虑最大值.

定理 3 Snell 包络满足

$$U_k = \max\{Z_k, \max_{\tau \in \mathcal{S}_{k+1,T}} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k)\},$$

对所有的 $k = 0, \cdots, T$.

证明 首先, 因为 $\tau = k$ 是一个 (常数) 停时, 又 $\mathcal{S}_{k+1,T} \subseteq \mathcal{S}_{k,T}$, 所以很清楚地有

$$\begin{aligned} U_k &= \max_{\tau \in \mathcal{S}_{k,T}} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k) \\ &\geqslant \max\{\mathcal{E}_\Pi(Z_k|\mathcal{P}_k), \max_{\tau \in \mathcal{S}_{k+1,T}} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k)\} \\ &= \max\{Z_k, \max_{\tau \in \mathcal{S}_{k+1,T}} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k)\}. \end{aligned}$$

我们必须建立相反的不等式. 令 $\tau \in \mathcal{S}_{k,T}$, 并考虑由 τ 变化而来的另一停时 τ', 由在 t_k 到 t_{k+1} 之间延迟任何停止得到. 也就是说,

$$\tau'(\omega) = \max\{\tau, k+1\} = \begin{cases} \tau(\omega), & \text{如果}\omega \in [\tau > k], \\ k+1, & \text{如果}\omega \in [\tau = k]. \end{cases}$$

因为两个停时的最大者也是一个停时, 所以有 $\tau' \in \mathcal{S}_{k+1,T}$.

现在, 由于 $[\tau > k] = [\tau = k]^c \in \sigma(\mathcal{P}_k)$, 因而有

$$\begin{aligned}\mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k) &= \mathcal{E}_\Pi(Z_\tau 1_{[\tau=k]}|\mathcal{P}_k) + \mathcal{E}_\Pi(Z_\tau 1_{[\tau>k]}|\mathcal{P}_k) \\ &= \mathcal{E}_\Pi(Z_k 1_{[\tau=k]}|\mathcal{P}_k) + \mathcal{E}_\Pi(Z_{\tau'} 1_{[\tau>k]}|\mathcal{P}_k) \\ &\leqslant \max\{Z_k, \mathcal{E}_\Pi(Z_{\tau'}|\mathcal{P}_k)\} \\ &= \max\{Z_k, \max_{\tau\in\ \mathcal{S}_{k+1,T}} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k)\}.\end{aligned}$$

但是上式左边对所有的 $\tau \in \mathcal{S}_{k,T}$ 都成立, 故

$$U_k = \max_{\tau\in\ \mathcal{S}_{k,T}} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k) \leqslant \max\{Z_k, \max_{\tau\in\ \mathcal{S}_{k+1,T}} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k)\}.$$

得证.

上述公式最重要的一个用处就是从该公式我们能推出 U_k 的一个向后循环关系. 注意到

$$U_T = \max_{\tau\in\mathcal{S}_{T,T}} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_T) = \mathcal{E}_\Pi(Z_T|\mathcal{P}_T) = Z_T.$$

这为向后的循环提供了第一步.

我们进一步来看看随机变量

$$X = \max_{\tau\in\mathcal{S}_{k+1,T}} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k),$$

这在定理 3 中出现过. 如果条件是关于 $\mathcal{P}_{k+1}$, 那么 X 就刚好是 U_{k+1}. 这就促使我们更深一层地考虑和利用条件期望的平滑性质. 首先, 需要提及的是, 在一般情形下, 对任意两个随机变量 X 和 Y, 有

$$\max\{\mathcal{E}(X|\mathcal{P}), \mathcal{E}(Y|\mathcal{P})\} \leqslant \mathcal{E}(\max\{X,Y\}|\mathcal{P}).$$

我们将证明留给读者. 现在有

$$\begin{aligned}X &= \max_{\tau\in\mathcal{S}_{k+1,T}} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k) \\ &= \max_{\tau\in\mathcal{S}_{k+1,T}} \mathcal{E}_\Pi(\mathcal{E}_\Pi(Z_\tau|\mathcal{P}_{k+1})|\mathcal{P}_k) \\ &\leqslant \mathcal{E}_\Pi(\max_{\tau\in\mathcal{S}_{k+1,T}} \{\mathcal{E}_\Pi(Z_\tau|\mathcal{P}_{k+1})\}|\mathcal{P}_k) \\ &= \mathcal{E}_\Pi(U_{k+1}|\mathcal{P}_k).\end{aligned}$$

故

$$X \leqslant \mathcal{E}_\Pi(U_{k+1}|\mathcal{P}_k).$$

对于反向不等式, 令 $\tau^* \in \mathcal{S}_{k+1,T}$ 是区间 $[t_{k+1}, T]$ 内的最优停时, 即

$$U_{k+1} = \mathcal{E}_\Pi(Z_{\tau^*}|\mathcal{P}_{k+1}).$$

那么

$$\begin{aligned}\mathcal{E}_\Pi(U_{k+1}|\mathcal{P}_k) &= \mathcal{E}_\Pi(\mathcal{E}_\Pi(Z_{\tau^*}|\mathcal{P}_{k+1})|\mathcal{P}_k)\\ &= \mathcal{E}_\Pi(Z_{\tau^*}|\mathcal{P}_k)\\ &\leqslant \max_{\tau\in\mathcal{S}_{k+1,T}} \mathcal{E}_\Pi(Z_\tau|\mathcal{P}_k)\\ &= X.\end{aligned}$$

因此

$$X = \mathcal{E}_\Pi(U_{k+1}|\mathcal{P}_k).$$

我们就得到了关于 Snell 包络的下述循环关系.

定理 4 Snell 包络满足向后循环关系

1) $U_T = Z_T$.

2)

$$U_k = \max\{Z_k, \mathcal{E}_\Pi(U_{k+1}|\mathcal{P}_k)\},$$

对所有的 $k = 0, \cdots, T-1$.

现在我们就能计算例 1 中损益过程的 Snell 包络了.

例 5 再考虑例 1, 我们来计算损益过程 (Y_k) 的 Snell 包络. 首先, 我们有

$$U_3 = Y_3.$$

接下来, 我们需要

$$\mathcal{E}_\Pi(U_3|\mathcal{P}_2) = \mathcal{E}_\Pi(Y_3|\mathcal{P}_2) = \begin{cases}\frac{1}{2}(5.26+0.78) = 3.2, & 如果\omega \in B_{2,1},\\ \frac{1}{2}(0.78+0) = 0.39, & 如果\omega \in B_{2,2},\\ \frac{1}{2}(0.78+0) = 0.39, & 如果\omega \in B_{2,3},\\ 0, & 如果\omega \in B_{2,4}.\end{cases}$$

从而得到

$$U_2 = \max\{Y_2, \mathcal{E}_\Pi(U_3|\mathcal{P}_2)\} = \begin{cases} 3.2, & \text{如果}\omega = \omega_1, \omega_2, \\ 0.39, & \text{如果}\omega = \omega_3, \omega_4, \\ 0.39, & \text{如果}\omega = \omega_5, \omega_6, \\ 0, & \text{如果}\omega = \omega_7, \omega_8. \end{cases}$$

接下来, 我们需要

$$\mathcal{E}_\Pi(U_2|\mathcal{P}_1) = \begin{cases} \frac{1}{2}(3.2 + 0.39) = 1.795, & \text{如果}\omega = \omega_1, \omega_2, \omega_3, \omega_4, \\ \frac{1}{2}(0.39 + 0) = 0.195, & \text{如果}\omega = \omega_5, \omega_6, \omega_7, \omega_8. \end{cases}$$

这就给出了

$$\begin{aligned} U_1 &= \max\{Y_1, \mathcal{E}_\Pi(U_2|\mathcal{P}_1)\} \\ &= \mathcal{E}_\Pi(U_2|\mathcal{P}_1) \\ &= \begin{cases} 1.795, & \text{如果}\omega = \omega_1, \omega_2, \omega_3, \omega_4, \\ 0.195, & \text{如果}\omega = \omega_5, \omega_6, \omega_7, \omega_8. \end{cases} \end{aligned}$$

最后

$$U_0 = \max\{Y_0, \mathcal{E}_\Pi(U_1|\mathcal{P}_0)\} = \max\left\{0, \frac{1}{2}(1.795 + 0.195)\right\} = 0.995.$$

回忆例 4 中定义的停时 σ,

$$\sigma(\omega) = \begin{cases} 2, \text{如果}\omega \in \{\omega_1, \omega_2, \omega_7, \omega_8\}, \\ 3, \text{其他}. \end{cases}$$

该停时带来的最终期望 (折现) 价值为

$$\mathcal{E}_\Pi(\overline{Y_\tau}) = \frac{1}{4} \cdot 3.2 + \frac{1}{4} \cdot 0.78 = 0.995.$$

这等于 U_0. 因此, σ 确实是一个最优停时. 事实上, 我们将看到 σ 是最小的最优停时, 意思是说它在任何其他最优停时之前停止 (通过观察知道, 我们可以等到 t_3 时的状态 ω_7 和 ω_8, 仍然可以达到最优).

现在正是时候来强调有关最优停时的一个问题, 即最优停时表示了在不能知道未来情形下该什么时候停止的最佳猜测. 因此, 一个最优停时不能保证产生最佳的可能损益. 确实, 从图 1 中明显看出最佳的可能执行程序包括在 t_2 时执行, 如果最终状态是 ω_2. 但如果最终状态是 ω_1 则需一直等到 t_3. 然而, 在 t_2 时我们不知道哪个状态将会出现: ω_1 还是 ω_2. 因此该计划不是一个停时.

最小的控制上鞅

从定理 4 的条件 2) 可以清楚地知道

$$\mathcal{E}_{\Pi}(U_{k+1}|\mathcal{P}_k) \leqslant U_k.$$

这正是 U_k 为上鞅的条件. 正式地, 称一个 $\mathbb{F}$ 适应的过程 (X_k) 是一个 $\mathbb{F}$ 上鞅, 如果

$$\mathcal{E}_{\Pi}(X_{k+1}|\mathcal{P}_k) \leqslant X_k.$$

也非常明显的是

$$Z_k \leqslant U_k.$$

即 U_k 控制了 Z_k. 利用循环关系不难看出 U_k 是控制 Z_k 的最小上鞅.

定理 5 Snell 包络 U_k 是控制 Z_k 的最小 $\mathbb{F}$ 上鞅.

证明 我们已经知道 U_k 是控制 Z_k 的一个上鞅. 假设 V_k 是控制 Z_k 的一个上鞅. 这等价于下面的简单不等式

$$V_k \geqslant \max\{Z_k, \mathcal{E}_{\Pi}(V_{k+1}|\mathcal{P}_k)\}.$$

现在可利用循环关系来逐步向后归纳. 在归纳中有基本的一步

$$V_T \geqslant Z_T = U_T.$$

假设 $V_{k+1} \geqslant U_{k+1}$, 则

$$V_k \geqslant \max\{Z_k, \mathcal{E}_{\Pi}(V_{k+1}|\mathcal{P}_k)\} \geqslant \max\{Z_k, \mathcal{E}_{\Pi}(U_{k+1}|\mathcal{P}_k)\} = U_k.$$

得证.

10.12 鞅的一些附加结果

为了使最优停时方面的讨论能够继续下去, 我们需要鞅和上鞅的一些其他结果.

定理 6 1) 如果 $\mathbb{X}$ 是一个 $\mathbb{F}$ 鞅, 那么对所有的 $j \leqslant k$, 有

$$\mathcal{E}(X_j) = \mathcal{E}(X_k).$$

2) 如果 $\mathbb{X}$ 是一个 $\mathbb{F}$ 上鞅, 那么对所有的 $j \leqslant k$, 有

$$\mathcal{E}(X_j) \geqslant \mathcal{E}(X_k).$$

证明 关于鞅有

$$\mathcal{E}(X_k|\mathcal{P}_j) = X_j.$$

取期望并利用平滑性得

$$\mathcal{E}(X_k) = \mathcal{E}(\mathcal{E}(X_k|\mathcal{P}_j)) = \mathcal{E}(X_j).$$

关于上鞅, 其证明类似.

停止–过程: Doob 选样停止定理

我们从随机过程的样本路径的正式定义开始.

定义 5 考虑一个随机过程 $\mathbb{X} = (X_0, \cdots, X_T)$. 对每个 $\omega \in \Omega$, 称序列

$$X_0(\omega), \cdots, X_T(\omega)$$

为一条样本路径.

从直观上看, 为了停止一随机过程, 当 τ 告诉我们这样去做的时候, 即在时刻 $\tau(\omega)$ 我们要停止每个 $\omega \in \Omega$ 对应的样本路径. 因此, 样本路径看起来就像

$$X_0(\omega), X_1(\omega), \cdots, X_{\tau(\omega)}(\omega), X_{\tau(\omega)}(\omega), \cdots.$$

故该路径的指标数等于 $n \wedge \tau(\omega) = \min\{n, \tau(\omega)\}$, 且可以写成

$$X_{0\wedge\tau(\omega)}(\omega), X_{1\wedge\tau(\omega)}(\omega), \cdots, X_{n\wedge\tau(\omega)}(\omega), X_{(n+1)\wedge\tau(\omega)}(\omega), \cdots.$$

定义 6 设 $\mathbb{X} = (X_0, \cdots, X_T)$ 是 $\mathbb{F}$ 适应的随机过程, τ 是 $\mathbb{F}$ 上的停时. 停止过程或者样本过程 (sampled process) 的定义为

$$\mathbb{X}^\tau = (X_\kappa^\tau) = (X_{k\wedge\tau}) = (X_{k\wedge 0}, \cdots, X_{k\wedge T}).$$

注意前三个表达式只是第四个表达式的记号.

观察到对每个 n 有

$$X_{n\wedge\tau} = \sum_{i=1}^{n-1} X_i 1_{\{\tau=i\}} + X_n 1_{\{\tau\geqslant n\}}. \tag{$*$}$$

下一定理是鞅论中最重要的结果之一. 停时也称为可选随机变量.

定理 7 (Doob 选样定理) 设 $\mathbb{X} = (X_k)$ 是一个鞅 (或者上鞅), τ 是一个停时. 那么停止过程 $\mathbb{X}^\tau$ 也是一个鞅 (或者上鞅).

证明 我们知道

$$\mathcal{E}(X_n|\mathcal{F}_{n-1}) = X_{n-1}.$$

据 $(*)$ 式有

$$\mathcal{E}(X_{n\wedge\tau}|\mathcal{F}_{n-1}) = \sum_{i=1}^{n-1}\mathcal{E}(X_i1_{\{\tau=i\}}|\mathcal{F}_{n-1}) + \mathcal{E}(X_n1_{\{\tau\geqslant n\}}|\mathcal{F}_{n-1}).$$

现在, 为了使式子简化, 因为对 $i \leqslant n-1$ 有 $\{\tau=i\}\in\mathcal{A}(\mathcal{P}_{n-1})$, 从而 X_i 是 $\mathcal{P}_{n-1}$ 可测的且 $\{\tau\geqslant n\}\in\mathcal{A}(\mathcal{P}_{n-1})$, 从而

$$\begin{aligned}\mathcal{E}(X_{n\wedge\tau}|\mathcal{F}_{n-1}) &= \sum_{i=1}^{n-1}X_i1_{\{\tau=i\}} + 1_{\{\tau\geqslant n\}}\mathcal{E}(X_n|\mathcal{F}_{n-1})\\ &= \sum_{i=1}^{n-1}X_i1_{\{\tau=i\}} + 1_{\{\tau\geqslant n\}}X_{n-1}\\ &= \sum_{i=1}^{n-2}X_i1_{\{\tau=i\}} + 1_{\{\tau=n-1\}}X_{n-1} + 1_{\{\tau\geqslant n\}}X_{n-1}\\ &= \sum_{i=1}^{n-2}X_i1_{\{\tau=i\}} + 1_{\{\tau\geqslant n-1\}}X_{n-1}\\ &= X_{(n-1)\wedge\tau}.\end{aligned}$$

这就是所要证的. 上鞅情形下的证明几乎一样.

Doob 分解

最后, 我们需要下面的分解定理.

定理 8 (Doob 分解定理) 设 $\mathbb{X}=(X_0,\cdots,X_T)$ 是一个 $\mathbb{F}$ 适应的随机过程.

1) 存在唯一的鞅 $\mathbb{M}=(M_0,\cdots,M_T)$ 和唯一的可料过程 $\mathbb{A}=(A_1,\cdots,A_T)$, 使得

$$X_k = M_k - A_k,$$

其中 $A_0=0$.

2) 如果 $\mathbb{X}$ 是一个上鞅, 那么 $\mathbb{A}$ 是非降的, 即 $A_{k+1}\geqslant A_k$.

证明 1), 令 $M_0=X_0, A_0=0$. 当 $k>0$ 时, 设

$$M_k = \sum_{i=1}^{k}[X_i - \mathcal{E}_\Pi(X_i|\mathcal{F}_{i-1})] + X_0,$$

且 $A_k = M_k - X_k$. 然后很容易验证这样构造的 $\mathbb{M},\mathbb{A}$ 满足要求.

2), 假设 $\mathbb{X}$ 是上鞅, 那么

$$\begin{aligned}M_k - A_k = X_k &\geqslant \mathcal{E}(X_{k+1}|\mathcal{F}_k)\\ &= \mathcal{E}(M_{k+1}|\mathcal{F}_k) - \mathcal{E}(A_{k+1}|\mathcal{F}_k)\\ &= M_k - A_{k+1}.\end{aligned}$$

所以 $A_k \leqslant A_{k+1}$.

10.13 最优停时的刻画

注意到如果两个离散随机变量满足 $X \leqslant Y$ 和 $\mathcal{E}_\Pi(X) = \mathcal{E}_\Pi(Y)$, 又由于 Π 是强正, 那么就有 $X = Y$. 我们将看到这个结果很有用.

结合前面学过的有关鞅和上鞅的一些结果, 我们就能转向手头的事情了, 即最优停时的刻画.

我们的出发点是, 如果我们停止随机过程 (U_k), 那么会发生什么? 因为 U_k 是上鞅, 所以由 Doob 选样定理可知对任意停时 $\tau \in \mathcal{S}_{0,T}$, 有停止过程 (U_k^τ) 也是一个上鞅, 即

$$\mathcal{E}_\Pi(U_{k+1}^\tau | \mathcal{P}_k) \leqslant U_k^\tau,$$

而且, 因为 $Z_k \leqslant U_k$, 故最终值满足 $Z_\tau \leqslant U_\tau$.

U_k^τ 是上鞅及 $Z_\tau \leqslant U_\tau$ 这个事实暗示着有不等式链

$$\mathcal{E}_\Pi(Z_\tau) \leqslant \mathcal{E}_\Pi(U_\tau) = \mathcal{E}_\Pi(U_T^\tau) \leqslant \cdots \leqslant \mathcal{E}_\Pi(U_k^\tau) \leqslant \cdots \leqslant \mathcal{E}_\Pi(U_0^\tau) = U_0.$$

现在, 如果 τ^* 是 $[0,T]$ 上的一个最优停时, 即

$$U_0 = \mathcal{E}_\Pi(Z_{\tau^*}).$$

那么前面的不等式序列就变成了等式序列. 因为 $Z_{\tau^*} \leqslant U_{\tau^*}$ 和 Π 是强正的, 从而这些等式中的第一个

$$\mathcal{E}_\Pi(Z_{\tau^*}) = \mathcal{E}_\Pi(U_{\tau^*})$$

暗示了 $Z_{\tau^*} = U_{\tau^*}$ (可看该节开始的备注).

进一步考察不等式链, 我们也看到

$$\mathcal{E}_\Pi(U_k^{\tau^*}) = \mathcal{E}_\Pi(U_{k-1}^{\tau^*}).$$

上鞅的性质表明 $(U_k^{\tau^*})$ 实际上是一个鞅. 为了看清这点, 上鞅的性质为

$$\mathcal{E}_\Pi(U_k^{\tau^*} | \mathcal{P}_{k-1}) \leqslant U_{k-1}^{\tau^*}.$$

但是两边有相同的期望值. 事实上, 左边取期望再根据平滑性就有

$$\mathcal{E}_\Pi(\mathcal{E}_\Pi(U_k^{\tau^*} | \mathcal{P}_{k-1})) = \mathcal{E}_\Pi(U_k^{\tau^*}) = \mathcal{E}_\Pi(U_{k-1}^{\tau^*}).$$

一般地, 如果 $A \leqslant B$ 且 $\mathcal{E}(A) = \mathcal{E}(B)$, 那么 $A = B$. 因而我们推出

$$\mathcal{E}_\Pi(U_k^{\tau^*} | \mathcal{P}_{k-1}) = U_{k-1}^{\tau^*}.$$

也就是说, $(U_k^{\tau^*})$ 是一个鞅.

关于另一方面, 假设 $Z_\tau = U_\tau$ 并且 (U_k^τ) 是一个鞅. 那么不等式序列变成等式序列, 特别有

$$U_0 = \mathcal{E}_\Pi(Z_\tau).$$

这表明 τ 是区间 $[0,T]$ 上的最优停时.

现在我们已经得到了最优停时的一个刻画.

定理 9 一个停时 $\tau \in \mathcal{S}_{0,T}$ 在区间 $[0,T]$ 上是最优的当且仅当

1) $Z_\tau = U_\tau$.

2) (U_k^τ) 是一个鞅.

10.14 最优停时和 Doob 分解

我们已经看到, 停时 τ 是最优的当且仅当 $Z_\tau = U_\tau$ 且 $\mathbb{U}^\tau = (U_k^\tau)$ 是一个鞅. 这就促使我们进一步看看 $\mathbb{U}^\tau$ 什么时候是鞅.

我们知道 Snell 包络

$$\mathbb{U} = (U_0, \cdots, U_T)$$

是一个上鞅. 利用 Doob 分解, 可以写成

$$\mathbb{U} = \mathbb{M} - \mathbb{A},$$

其中 $\mathbb{M} = (M_0, \cdots, M_T)$ 是一个鞅, $\mathbb{A} = (A_1, \cdots, A_T)$ 是可料的且非降, $A_0 = 0$.

假设我们现在停止该序列

$$\mathbb{U}^\tau = \mathbb{M}^\tau - \mathbb{A}^\tau = (M_0^\tau - A_0^\tau, \cdots, M_T^\tau - A_T^\tau).$$

我们知道 $\mathbb{M}^\tau$ 是一个鞅. 从 Doob 分解的唯一性可知 $\mathbb{U}^\tau$ 是一个鞅当且仅当 $\mathbb{A}^\tau$ 是个零过程.

现在, 对任何 $\omega \in \Omega$, 序列 $A_k^\tau(\omega)$ 为

$$A_0(\omega), \cdots, A_{\tau(\omega)-1}(\omega), A_{\tau(\omega)}(\omega), \cdots, A_{\tau(\omega)}(\omega).$$

既然该序列从 0 开始是非降的, 那么它是零序列当且仅当

$$[A_\tau](\omega) = A_{\tau(\omega)}(\omega) = 0.$$

因此, $\mathbb{A}^\tau = 0$ 当且仅当 $A_\tau = 0$, 也就是说, U^τ 是一个鞅当且仅当 $A_\tau = 0$.

定理 10 设 $\mathbb{U} = (U_k)$ 是 (Z_k) 的 Snell 包络. 对一个停时 $\tau \in \mathcal{S}_{0,T}$, 停止过程 $\mathbb{U}^\tau = (U_k^\tau)$ 是一个鞅当且仅当 $A_\tau = 0$, 这里 $\mathbb{A} = (A_k)$ 是 $\mathbb{U}$ 的 Doob 分解中的可料过程.

10.15　最小的最优停时

前面的定理使得确定最小的最优停时变得容易起来. 首先, 我们回顾一下循环公式

$$U_k = \max\{Z_k, \mathcal{E}_\Pi(U_{k+1}|\mathcal{P}_k)\}.$$

利用 Doob 分解, 我们注意到

$$\begin{aligned}\mathcal{E}_\Pi(U_{k+1}|\mathcal{P}_k) &= \mathcal{E}_\Pi(M_{k+1}|\mathcal{P}_k) - \mathcal{E}_\Pi(A_{k+1}|\mathcal{P}_k)\\ &= M_k - A_{k+1}\\ &= (M_k - A_k) - (A_{k+1} - A_k)\\ &= U_k - (A_{k+1} - A_k).\end{aligned}$$

所以

$$U_k = \max\{Z_k, U_k - (A_{k+1} - A_k)\}.$$

因此, 严格不等式

$$U_k > Z_k$$

暗示着

$$A_{k+1} = A_k.$$

这导致在时刻 $t_k(U_k = Z_k)$ 之前有严格不等式 $U_i > Z_i (i < k)$, 故 $A_{i+1} = A_i (i < k)$. 但是 $A_0 = 0$, 因此

$$0 = A_0 = \cdots = A_k.$$

这就促使我们定义 $\tau_{\min}$ 如下

$$\tau_{\min}(\omega) = \min\{k | Z_k(\omega) = U_k(\omega)\}.$$

这样的 $\tau_{\min}$ 存在, 因为 $Z_T(\omega) = U_T(\omega)$. 另外, $\tau_{\min}$ 是适应过程 $(Z_k - U_k)$ 首次进入 Borel 集 $\{0\}$ 的时间, 因此它是一个停时. 由定义有

$$Z_{\tau_{\min}} = U_{\tau_{\min}}.$$

如果 $\tau_{\min}$ 是一个最优停时, 那么它一定是最小的最优停时, 因为所有的最优停时 τ 满足 $Z_\tau = U_\tau$, 即 $Z_{\tau_\omega}(\omega) = U_{\tau_\omega}(\omega)$. 而且, 我们刚刚知道

$$0 = A_0(\omega) = A_1(\omega) = \cdots = A_{\tau_{\min}(\omega)}(\omega).$$

因此 $A_{\tau_{\min}} = 0$, 这表明 $\mathbb{U}^{\tau_{\min}}$ 是一个鞅.

定理 11　最小的最优停时为

$$\tau_{\min}(\omega) \in \min\{k | Z_k(\omega) = U_k(\omega)\}.$$

例 6　在例 4 中我们定义停时

$$\sigma(\omega) = \begin{cases} 2, & 如果\omega \in \{\omega_1, \omega_2, \omega_7, \omega_8\}, \\ 3, & 其他. \end{cases}$$

在例 5 中我们指出了 σ 是最优的. 事实上, 很容易知道 σ 具有性质

$$\sigma(\omega) \in \min\{k | Z_k(\omega) = U_k(\omega)\}.$$

因此它是最小的最优停时.

10.16　最大的最优停时

考虑到定理 10, 在谈论最大的最优停时很自然考虑函数

$$\tau_{\max}(\omega) = \max\{k | A_k(\omega) = 0\}.$$

因为 $A_0 = 0$, 故 $\tau_{\max}$ 存在. 既然任何最优停时 τ 都满足 $A_\tau = 0$, 如果 $\tau_{\max}$ 是一个最优停时, 那么它一定是最大的最优停时. 注意, $\tau_{\max}$ 也能定义如下

$$\tau_{\max}(\omega) = \begin{cases} \min\{k | A_{k+1}(\omega) > 0\}, & 如果\{k | A_{k+1}(\omega) > 0\} \neq \varnothing, \\ T, & 其他情况. \end{cases}$$

又因为它是适应过程 $\mathbb{A}$ 首次进入 $(0, \infty)$ 的时间, 所以 $\tau_{\max}$ 是一个停时. 又 $A_{\tau_{\max}} = 0$, 根据定理 10 我们知道 $(U_k^{\tau_{\max}})$ 是一个鞅. 因此, 为了证明 $\tau_{\max}$ 是最优的, 只需证明 $U_{\tau_{\max}} = Z_{\tau_{\max}}$.

我们又一次来看循环关系

$$U_k = \max\{Z_k, \mathcal{E}_\Pi(U_{k+1} | \mathcal{P}_k)\}.$$

利用 Doob 分解, 可写成

$$U_k = \max\{Z_k, U_k - (A_{k+1} - A_k)\}.$$

但是当 $k = \tau_{\max}(\omega)$ 时, 有

$$A_{\tau_{\max}+1}(\omega) - A_{\tau_{\max}}(\omega) = A_{\tau_{\max}+1}(\omega) > 0.$$

所以上述最大值刚好是 Z_k, 即

$$U_{\tau_{\max}}(\omega) = Z_{\tau_{\max}}(\omega).$$

因此 $U_{\tau_{\max}} = Z_{\tau_{\max}}$, 这就是所要的.

定理 12　最大的最优停时是

$$\begin{aligned}\tau_{\max}(\omega) &= \max\{k|A_k(\omega)=0\}\\ &= \begin{cases}\min\{k|A_{k+1}(\omega)>0\}, & \text{如果}\{k|A_{k+1}(\omega)>0\}\neq\varnothing,\\ T, & \text{其他情况},\end{cases}\end{aligned}$$

这里 $U_k = M_k - A_k$ 是 Doob 分解.

练　习　10

1. 证明：形如 (a,b) 这样的集合的首次进入时是一个停时.

2. 证明：任何 Borel 集的首次进入时是一个停时.

3. 证明：股票的价格达到其初始价格两倍的首次达到时是一个停时.

4. 证明：股票的价格达到其先前价格的两倍的首次达到时是一个停时, 也就是说, 下面定义的随机变量是一个停时

$$\tau(\omega) = \begin{cases}\min\{k|S_k(\omega)\geqslant 2S_{k-1}(\omega)\}, & \text{如果}\{k|S_k(\omega)\geqslant 2S_{k-1}(\omega)\}\neq\varnothing,\\ T, & \text{其他情况}.\end{cases}$$

5. 证明：形如 (a,b) 这样的集合的首次离开时是一个停时.

6. 证明：两个停时的最大值、最小值或者之和仍是一个停时. 之差呢?

7. 证明：对任何随机变量 X 和 Y, 有

$$\max\{\mathcal{E}(X|\mathcal{P}),\mathcal{E}(Y|\mathcal{P})\}\leqslant\mathcal{E}(\max\{X,Y\}|\mathcal{P}).$$

8. 证明：如果对所有的 k 有 $Z_k\leqslant U_k$, 那么对任何停时 τ 都有 $Z_\tau\leqslant U_\tau$.

9. 证明：如果两个离散随机变量满足 $X\leqslant Y$ 和 $\mathcal{E}_\Pi(X)=\mathcal{E}_\Pi(Y)$, 那么因为 Π 是强正, 所以就有 $X=Y$.

10. 证明：如果 $(A_0,\cdots,A_T)$ 是一个鞅且 $(A_1,\cdots,A_T)$ 是可料的, 那么 (A_k) 是一个常数序列, 即

$$A_0^\tau=\cdots=A_T^\tau.$$

11. 在 CRR 模型中

$$u=1.2,\quad d=0.8,\quad r=0,\quad S_0=20.$$

一个美式期权的执行价为 $K = 21$, 请计算其价格过程, 损益过程和 Snell 包络. 找出最小的最优停时.

12. 考虑给定的 u, d, r, S_0 和 K, 制作一张 Excel 表用来计算执行价为 K 的美式买权/卖权的价格过程, 损益过程和 Snell 包络.

参考答案节选

练　习　1

1. 1/2.

3. 5/12, 9/12, 7/12.

5. 1/3, 1/6,1/3, 1/2.

7. 11/16.

11. 1/2.

13. 25/13 分. 是.

21. 假设 X 的取值为 $\{x_1,\cdots,x_n\}$, Y 的取值为 $\{y_1,\cdots,y_m\}$. 那么 XY 的取值为不同的乘积 x_iy_j. 对 XY 的一个给定取值 a, 令

$$\{(x_{i_1},y_{j_1}),\cdots,(x_{i_m},y_{j_m})\}$$

为乘积为 a 的所有序对所形成的集合. 那么

$$\begin{aligned}
\mathbb{P}(XY=a,Z=z)&=\mathbb{P}\left\{\left[\bigcup_{k=1}^{m}(\{X=x_{i_k}\}\cap\{Y=y_{j_k}\})\right]\cap\{Z=z\}\right\}\\
&=\mathbb{P}\left\{\bigcup_{k=1}^{m}(\{X=x_{i_k}\}\cap\{Y=y_{j_k}\}\cap\{Z=z\})\right\}\\
&=\sum_{k=1}^{m}\mathbb{P}(X=x_{i_k})\mathbb{P}(Y=y_{j_k})\mathbb{P}(Z=z)\\
&=\left[\sum_{k=1}^{m}\mathbb{P}(X=x_{i_k})\mathbb{P}(Y=y_{j_k})\right]\mathbb{P}(Z=z)\\
&=\left[\sum_{k=1}^{m}\mathbb{P}(X=x_{i_k},Y=y_{j_k})\right]\mathbb{P}(Z=z)\\
&=\mathbb{P}(XY=a,Z=z).
\end{aligned}$$

练　习　2

1. $\beta=1$.

3. 通过解 σ' 的方程得到 s 关于 σ' 的线性函数, 然后将 s 代入 μ 的方程即可得到 μ 为 σ' 的线性函数.

6. 我们有

$$
\begin{aligned}
\operatorname{cov}(R_M,\varepsilon) &= \operatorname{cov}(R_M, R_i-\beta_k R_M) \\
&= \operatorname{cov}(R_M,R_i)-\beta_k\operatorname{cov}(R_M,R_M) \\
&= \operatorname{cov}(R_M,R_i)-\frac{\operatorname{cov}(R_i,R_M)}{\operatorname{cov}(R_M,R_M)}\operatorname{cov}(R_M,R_M) \\
&= 0.
\end{aligned}
$$

7. 从资本市场线方程可求解出 β, 然后将 β 代入线性回归方程 $y=\beta x+\alpha$, 得到方程

$$
y=\frac{\mu_k-\mu_{rf}}{\mu_M-\mu_{rf}}(x-\mu_M)+\mu_k.
$$

令 $x=\mu_{rf}$, 得 $y=\mu_{rf}$

练　习　3

4. 执行价低的买权的成本 C_1 高于执行价高的买权的成本 C_2. 利润曲线如图 1 所示.

7. 利润曲线如图 2 所示.

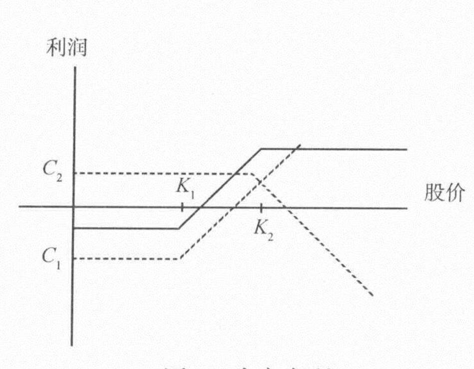


图 1　牛市套利

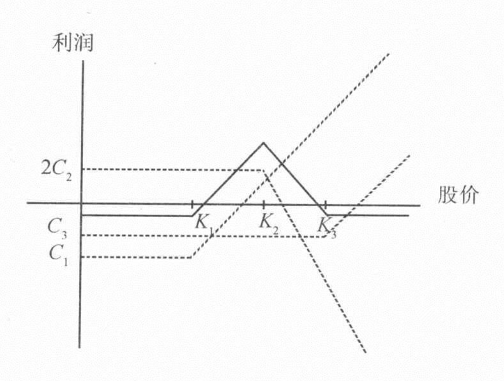


图 2　蝶式套利

练　习　4

1. 我们仍然有

$$
\mathcal{V}(\text{多头合约})=S_T-F_{0,T},
$$
$$
\mathcal{V}(\text{空头合约})=F_{0,T}-S_T.
$$

但是正向 – 买进持有投资者有最终损益

$$\mathcal{V}(\text{正向 – 买现卖期}) = S_T - S_0 e^{rT} + Ie^{rT}$$

和

$$\mathcal{V}(\text{反向 – 买现卖期}) = S_0 e^{rT} - S_T - Ie^{rT}.$$

为了解释最后一项, 注意一种资产的卖空要求有贷方借出该资产. 贷方不仅要求获得资产本身的收益, 而且要求获得由于拥有资产带来的好处这部分收入.

3. 令练习 2 中的最终损益为 0, 得到

$$F_{0,T} = (S_0 - I)e^{rT}.$$

5. 策略 1 的最终损益为

$$\mathcal{V}(\text{多头合约}) + \mathcal{V}(\text{反向 – 买现卖期}) = S_0 e^{r_l T} - F_{0,T}.$$

策略 2 的最终损益为

$$\mathcal{V}(\text{空头合约}) + \mathcal{V}(\text{正向 – 买现卖期}) = F_{0,T} - S_0 e^{r_b T}.$$

7. 如果不是, 买入股票卖出买权, 将获得差额利润. 如果买权被执行, 用持有的股票对冲即可.

9. 利用卖权 – 买权平价公式.

练　习　5

3. 我们有

$$\begin{aligned}\mathcal{E}(X|\mathcal{P})(\omega) &= \mathcal{E}(X|[\omega]_{\mathcal{P}})\\ &= \sum_{\sigma\in\Omega} X(\sigma)\mathbb{P}(\sigma|[\omega]_{\mathcal{P}})\\ &\geqslant 0.\end{aligned}$$

17. 对 $\sigma \in \Omega$, 令 $k = N(\omega)$. 那么

$$\begin{aligned}\mathcal{E}(S|N)(\omega) &= \mathcal{E}(S|N=k)\\ &= \sum_{i=1}^{m} r_i \mathbb{P}(S=r_i|N=k)\\ &= \sum_{i=1}^{m} r_i \frac{\mathbb{P}((X_1+\cdots+X_N=r_i)\cap(N=k))}{\mathbb{P}(N=k)}\end{aligned}$$

$$
\begin{aligned}
&=\sum_{i=1}^{m} r_i \frac{\mathbb{P}(X_1+\cdots+X_k=r_i)\mathbb{P}(N=k)}{\mathbb{P}(N=k)}\\
&=\sum_{i=1}^{m} r_i \mathbb{P}(X_1+\cdots+X_k=r_i)\\
&=\mathcal{E}(X_1+\cdots+X_k)\\
&=\mu k\\
&=\mu N(\omega).
\end{aligned}
$$

18. 其解为

$$
\begin{aligned}
\mathcal{E}(\mathcal{E}(X|\mathcal{Q})|\mathcal{P})&=\mathcal{E}\left(\sum_{i=1}^{m}\mathcal{E}(X|C_i)1_{C_i}|\mathcal{P}\right)\\
&=\sum_{i=1}^{m}\mathcal{E}(\mathcal{E}(X|C_i)1_{C_i}|\mathcal{P})\\
&=\sum_{i=1}^{m}\mathcal{E}(X|C_i)\mathcal{E}(1_{C_i}|\mathcal{P})\\
&=\sum_{i=1}^{m}\mathcal{E}(X|C_i)\left[\sum_{j=1}^{k}\mathcal{E}(1_{C_i}|B_j)1_{B_j}\right]\\
&=\sum_{i=1}^{m}\mathcal{E}(X|C_i)\left[\sum_{j=1}^{k}\frac{\mathcal{E}(1_{C_i}1_{B_j})}{\mathbb{P}(B_j)}1_{B_j}\right]\\
&=\sum_{i=1}^{m}\mathcal{E}(X|C_i)\left[\sum_{j=1}^{k}\frac{\mathcal{E}(1_{C_i\cup B_j})}{\mathbb{P}(B_j)}1_{B_j}\right].
\end{aligned}
$$

现在, 因为对每个 C_i 存在唯一的 B_{j_i}, 使得 $C_i\subseteq B_{j_i}$, 所以我们知道

$$
1_{C_i\cup B_j}=\begin{cases}1_{C_i}, & j=j_i,\\ 0, & j\neq j_i.\end{cases}
$$

故认识到

$$
\begin{aligned}
\sum_{i=1}^{m}\mathcal{E}(X|C_i)\left[\sum_{j=1}^{k}\frac{\mathcal{E}(1_{C_i\cup B_j})}{\mathbb{P}(B_j)}1_{B_j}\right]&=\sum_{i=1}^{m}\mathcal{E}(X|C_i)\frac{\mathcal{E}(1_{C_i})}{\mathbb{P}(B_{j_i})}1_{B_{j_i}}\\
&=\sum_{i=1}^{m}\frac{\mathcal{E}(X1_{C_i})}{\mathbb{P}(C_i)}\frac{\mathcal{E}(1_{C_i})}{\mathbb{P}(B_{j_i})}1_{B_{j_i}}\\
&=\sum_{i=1}^{m}\frac{\mathcal{E}(X1_{C_i})}{\mathbb{P}(B_{j_i})}1_{B_{j_i}}.
\end{aligned}
$$

现在将求和部分分成更小的求和, 将包含在块 B_{j_i} 中的 C_i 合在一起 (换句话说, 按 $\mathcal{P}$ 的块来分组求和). 这就得到

$$\sum_{i=1}^{m} \frac{\mathcal{E}(X1_{C_i})}{\mathbb{P}(B_{j_i})} 1_{B_{j_i}} = \sum_{j=1}^{k} \frac{\mathcal{E}(X1_{B_j})}{\mathbb{P}(B_j)} 1_{B_j} = \mathcal{E}(X|\mathcal{P}).$$

这就是我们所要的.

练 习 6

1. 方程系统为

$$\begin{aligned}
\mathcal{V}_2(\Theta_2)(\omega_1) &= 95, \\
\mathcal{V}_2(\Theta_2)(\omega_2) &= 90, \\
\mathcal{V}_2(\Theta_2)(\omega_3) &= 85, \\
\mathcal{V}_2(\Theta_2)(\omega_4) &= 75,
\end{aligned}$$

或

$$\begin{aligned}
S_{2,1}(\omega_1)\theta_{2,1}(\omega_1) + S_{2,2}(\omega_1)\theta_{2,2}(\omega_1) &= 95, \\
S_{2,1}(\omega_2)\theta_{2,1}(\omega_2) + S_{2,2}(\omega_2)\theta_{2,2}(\omega_2) &= 90, \\
S_{2,1}(\omega_3)\theta_{2,1}(\omega_3) + S_{2,2}(\omega_3)\theta_{2,2}(\omega_3) &= 85, \\
S_{2,1}(\omega_4)\theta_{2,1}(\omega_4) + S_{2,2}(\omega_4)\theta_{2,2}(\omega_4) &= 75.
\end{aligned}$$

将实际价格代入得

$$\begin{aligned}
\theta_{2,1}(\omega_1) + 90\theta_{2,2}(\omega_1) &= 95, \\
\theta_{2,1}(\omega_2) + 80\theta_{2,2}(\omega_2) &= 90, \\
\theta_{2,1}(\omega_3) + 80\theta_{2,2}(\omega_3) &= 85, \\
\theta_{2,1}(\omega_4) + 75\theta_{2,2}(\omega_4) &= 75.
\end{aligned}$$

利用 Θ_2 是 $\mathcal{P}_1$ 可测的, 有

$$\begin{aligned}
\theta_{2,1}(\omega_1) &= \theta_{2,1}(\omega_2), \\
\theta_{2,1}(\omega_3) &= \theta_{2,1}(\omega_4), \\
\theta_{2,2}(\omega_1) &= \theta_{2,2}(\omega_2), \\
\theta_{2,2}(\omega_3) &= \theta_{2,2}(\omega_4).
\end{aligned}$$

因此前面的方程组系统可仅用 ω_1, ω_3 来表示

$$\begin{aligned} \theta_{2,1}(\omega_1) + 90\theta_{2,2}(\omega_1) &= 95, \\ \theta_{2,1}(\omega_1) + 80\theta_{2,2}(\omega_1) &= 90, \\ \theta_{2,1}(\omega_3) + 80\theta_{2,2}(\omega_3) &= 85, \\ \theta_{2,1}(\omega_3) + 75\theta_{2,2}(\omega_3) &= 75. \end{aligned}$$

前面两个方程有唯一解, 后面两个方程也一样, 解之得

$$\begin{aligned} \Theta_2(\omega_1) = \Theta_2(\omega_2) &= \left(50, \frac{1}{2}\right), \\ \Theta_2(\omega_3) = \Theta_2(\omega_4) &= (-75, 2). \end{aligned}$$

根据时间往后倒推, 我们接下来计算 Θ_2 的占有价值

$$\begin{aligned} \mathcal{V}_1(\Theta_2)(\omega_1) &= 50 + 85 \cdot \frac{1}{2} = \frac{185}{2}, \\ \mathcal{V}_1(\Theta_2)(\omega_3) &= -75 + 78 \cdot 2 = 81. \end{aligned}$$

自融资条件要求 Θ_1 的清算价值等于 Θ_2 的占有价值, 因此

$$\begin{aligned} \mathcal{V}_1(\Theta_1)(\omega_1) &= \frac{185}{2}, \\ \mathcal{V}_1(\Theta_1)(\omega_3) &= 81. \end{aligned}$$

将上述方程展开并代入实际价格, 得

$$\begin{aligned} \theta_{1,1}(\omega_1) + 85\theta_{1,2}(\omega_1) &= \frac{185}{2}, \\ \theta_{1,1}(\omega_3) + 78\theta_{1,2}(\omega_3) &= 81. \end{aligned}$$

但是 Θ_1 是 $\mathcal{P}_0$ 可测的, 即 Θ_1 在 Ω 上是常数, 因此对任意 $\omega \in \Omega$, 有

$$\begin{aligned} \theta_{1,1}(\omega) + 85\theta_{1,2}(\omega) &= \frac{185}{2}, \\ \theta_{1,1}(\omega) + 78\theta_{1,2}(\omega) &= 81. \end{aligned}$$

该方程组有解

$$\Theta_1(\omega) = \left(-\frac{330}{7}, \frac{23}{14}\right).$$

即投资组合由 330/7 份债券空头和 23/14 份股票多头组成, 其初始价值为

$$-\frac{330}{7} + 80 \cdot \frac{23}{14} \approx 84.29.$$

2. 方程系统为

$$\begin{aligned}
\mathcal{V}_3(\Theta_3)(\omega_1) &= 100,\\
\mathcal{V}_3(\Theta_3)(\omega_2) &= 100,\\
\mathcal{V}_3(\Theta_3)(\omega_3) &= 95,\\
\mathcal{V}_3(\Theta_3)(\omega_4) &= 90,\\
\mathcal{V}_3(\Theta_3)(\omega_5) &= 90,\\
\mathcal{V}_3(\Theta_3)(\omega_6) &= 85,
\end{aligned}$$

或者, 因为无风险资产 a_1 的价格为 1,

$$\begin{aligned}
\theta_{3,1}(\omega_1) + S_{3,2}(\omega_1)\theta_{3,2}(\omega_1) &= 100,\\
\theta_{3,1}(\omega_2) + S_{3,2}(\omega_2)\theta_{3,2}(\omega_2) &= 100,\\
\theta_{3,1}(\omega_3) + S_{3,2}(\omega_3)\theta_{3,2}(\omega_3) &= 95,\\
\theta_{3,1}(\omega_4) + S_{3,2}(\omega_4)\theta_{3,2}(\omega_4) &= 90,\\
\theta_{3,1}(\omega_5) + S_{3,2}(\omega_5)\theta_{3,2}(\omega_5) &= 90,\\
\theta_{3,1}(\omega_6) + S_{3,2}(\omega_6)\theta_{3,2}(\omega_6) &= 85.
\end{aligned}$$

将实际价格代入得

$$\begin{aligned}
\theta_{3,1}(\omega_1) + 100\theta_{3,2}(\omega_1) &= 100,\\
\theta_{3,1}(\omega_2) + 95\theta_{3,2}(\omega_2) &= 100,\\
\theta_{3,1}(\omega_3) + 95\theta_{3,2}(\omega_3) &= 95,\\
\theta_{3,1}(\omega_4) + 90\theta_{3,2}(\omega_4) &= 90,\\
\theta_{3,1}(\omega_5) + 90\theta_{3,2}(\omega_5) &= 90,\\
\theta_{3,1}(\omega_6) + 80\theta_{3,2}(\omega_6) &= 85.
\end{aligned}$$

利用 Θ_3 是 $\mathcal{P}_2$ 可测的, 有

$$\begin{aligned}
\theta_{3,1}(\omega_1) &= \theta_{3,1}(\omega_2),\\
\theta_{3,1}(\omega_4) &= \theta_{3,1}(\omega_5),\\
\theta_{3,2}(\omega_1) &= \theta_{3,2}(\omega_2),\\
\theta_{3,2}(\omega_4) &= \theta_{3,2}(\omega_5).
\end{aligned}$$

因此前面的方程组系统可仅用 $\omega_1, \omega_3, \omega_4, \omega_6$ 来表示

$$\begin{aligned}
\theta_{3,1}(\omega_1) + 100\theta_{3,2}(\omega_1) &= 100, \\
\theta_{3,1}(\omega_1) + 95\theta_{3,2}(\omega_1) &= 100, \\
\theta_{3,1}(\omega_3) + 95\theta_{3,2}(\omega_3) &= 95, \\
\theta_{3,1}(\omega_4) + 90\theta_{3,2}(\omega_4) &= 90, \\
\theta_{3,1}(\omega_4) + 90\theta_{3,2}(\omega_4) &= 90, \\
\theta_{3,1}(\omega_6) + 80\theta_{3,2}(\omega_6) &= 85.
\end{aligned}$$

前面两个方程有唯一解, 第四和第五两个方程也一样, 解之得

$$\begin{aligned}
\Theta_3(\omega_1) = \Theta_3(\omega_2) = (100, 0), \\
\Theta_3(\omega_4) = \Theta_3(\omega_5) = (0, 1).
\end{aligned}$$

同时

$$\begin{aligned}
\Theta_3(\omega_3) &= \left(s, \frac{95-s}{95}\right), \\
\Theta_3(\omega_6) &= \left(t, \frac{85-t}{80}\right),
\end{aligned}$$

这里 s 和 t 为参数. Θ_3 的占有价值为

$$\begin{aligned}
&\mathcal{V}_2(\Theta_3)(\omega_1) = \mathcal{V}_2(\Theta_3)(\omega_2) = 100, \\
&\mathcal{V}_2(\Theta_3)(\omega_3) = s + 80 \cdot \frac{95-s}{95} = \frac{3s}{19} + 80, \\
&\mathcal{V}_2(\Theta_3)(\omega_4) = \mathcal{V}_2(\Theta_3)(\omega_5) = 80, \\
&\mathcal{V}_2(\Theta_3)(\omega_6) = t + 75 \cdot \frac{95-t}{95} = \frac{4t}{19} + 75.
\end{aligned}$$

自融资条件要求上述值也是 Θ_2 的清算价值, 故 Θ_2 必须复制未定权益

$$\left(100, \frac{3s}{19} + 80, 80, \frac{4t}{19} + 75\right).$$

既然要求只有一个复制组合, 我们令 $s = t = 0$ 得未定权益为

$$(100, 80, 80, 75).$$

得到方程系统

$$\begin{aligned}
\theta_{2,1}(\omega_1)+90\theta_{2,2}(\omega_1)&=100,\\
\theta_{2,1}(\omega_2)+90\theta_{2,2}(\omega_2)&=100,\\
\theta_{2,1}(\omega_3)+80\theta_{2,2}(\omega_3)&=80,\\
\theta_{2,1}(\omega_4)+80\theta_{2,2}(\omega_4)&=80,\\
\theta_{2,1}(\omega_5)+80\theta_{2,2}(\omega_5)&=80,\\
\theta_{2,1}(\omega_6)+75\theta_{2,2}(\omega_6)&=75.
\end{aligned}$$

由于 $\theta_{2,i}$ 在块 $\{\omega_1,\omega_2,\omega_3\}$ 和块 $\{\omega_4,\omega_5,\omega_6\}$ 上为常数, 故可写成

$$\begin{aligned}
\theta_{2,1}(\omega_1)+90\theta_{2,2}(\omega_1)&=100,\\
\theta_{2,1}(\omega_1)+90\theta_{2,2}(\omega_1)&=100,\\
\theta_{2,1}(\omega_1)+80\theta_{2,2}(\omega_1)&=80,\\
\theta_{2,1}(\omega_4)+80\theta_{2,2}(\omega_4)&=80,\\
\theta_{2,1}(\omega_4)+80\theta_{2,2}(\omega_4)&=80,\\
\theta_{2,1}(\omega_4)+75\theta_{2,2}(\omega_4)&=75,
\end{aligned}$$

或者

$$\begin{aligned}
\theta_{2,1}(\omega_1)+90\theta_{2,2}(\omega_1)&=100,\\
\theta_{2,1}(\omega_1)+80\theta_{2,2}(\omega_1)&=80,\\
\theta_{2,1}(\omega_4)+80\theta_{2,2}(\omega_4)&=80,\\
\theta_{2,1}(\omega_4)+75\theta_{2,2}(\omega_4)&=75.
\end{aligned}$$

可得

$$\begin{aligned}
\Theta_2(\omega_1)=\Theta_2(\omega_2)=\Theta_2(\omega_3)&=(-80,2),\\
\Theta_2(\omega_4)=\Theta_2(\omega_5)=\Theta_2(\omega_6)&=(0,1).
\end{aligned}$$

根据时间往后倒推, 我们接下来计算 Θ_2 的占有价值

$$\begin{aligned}
\mathcal{V}_1(\Theta_2)(\omega_1)&=-80+85\cdot 2=90,\\
\mathcal{V}_1(\Theta_2)(\omega_4)&=0+78=78.
\end{aligned}$$

自融资条件要求上述值也是 Θ_1 的清算价值, 因此

$$\begin{aligned}
\mathcal{V}_1(\Theta_1)(\omega_1)&=90,\\
\mathcal{V}_1(\Theta_1)(\omega_4)&=78.
\end{aligned}$$

将上述方程展开并代入实际价格得到

$$\theta_{1,1}(\omega_1) + 85\theta_{1,2}(\omega_1) = 90,$$
$$\theta_{1,1}(\omega_4) + 78\theta_{1,2}(\omega_4) = 78.$$

但是 Θ_1 是 $\mathcal{P}_0$ 可测的, 即 Θ_1 在 Ω 上是常数, 因此对任意 $\omega \in \Omega$, 有

$$\theta_{1,1}(\omega) + 85\theta_{1,2}(\omega) = 90,$$
$$\theta_{1,1}(\omega) + 78\theta_{1,2}(\omega) = 78.$$

该方程组有解

$$\Theta_1(\omega) = \left(-\frac{390}{7}, \frac{12}{7}\right).$$

即投资组合由 390/7 份债券空头和 12/7 份股票多头组成, 其初始价值为

$$-\frac{390}{7} + 80 \cdot \frac{12}{7} = \frac{570}{7} \approx 81.43.$$

3. 第一枚为反面：庄家不亏不赢, 玩家输 \$1 百万, 赌博结束.

第一枚为正面：庄家赢 \$2 百万, 玩家输 \$1 百万, 赌博继续.

第二枚为反面：庄家不亏不赢, 玩家输 \$1 百万, 赌博结束.

第二枚为正面：庄家赢 \$4 百万, 玩家输 \$1 百万, 赌博继续.

第三枚为反面：庄家不亏不赢, 玩家输 \$1 百万, 赌博结束.

第三枚为正面：庄家不亏不赢, 玩家赢 \$8 百万 (包括本金), 游戏结束.

在所有结束的情形当中, 庄家都是不亏不赢. 因此, 庄家有一个完美的对冲. 因为庄家从来没投入自己的钱和拿走钱, 所以庄家的策略是自融资的. 庄家用与做市商赌博的损益复制了与玩家之间的损益, 这样使得庄家的损益为 0.

5. 自融资条件为：对所有的 $i = 1, \cdots, T-1$, 有

$$\mathcal{V}_i(\Theta_i') = \mathcal{V}_i(\Theta_{i+1}').$$

因为根据假设有 Φ 是自融资的, 所以 Θ_i' 的清算价值为

$$\begin{aligned}\mathcal{V}_i(\Theta_i') &= (\theta_{i,1} + a)S_{i,1} + \sum_{j=2}^{n} \theta_{i,j}S_{i,j} \\ &= \mathcal{V}_i(\Theta_i) + aS_{i,1}1_\Omega \\ &= \mathcal{V}_i(\Theta_{i+1}) + aS_{i,1}1_\Omega.\end{aligned}$$

而占有价值为

$$\begin{aligned}\mathcal{V}_i(\Theta'_{i+1}) &= (\theta_{i+1,1}+a)S_{i,1}+\sum_{j=2}^{n}\theta_{i+1,j}S_{i,j}\\ &= \mathcal{V}_i(\Theta_{i+1})+aS_{i,1}1_{\Omega}.\end{aligned}$$

因此 Φ' 是自融资的.

7. 令 $\Phi_0=0$ 是零交易策略 (这里所有的投资组合都是零组合). 如果一价定律成立, 那么对任意初始价值为零的交易策略 Φ, 有

$$\begin{aligned}\mathcal{V}_T(\Phi)=0 &\Rightarrow \mathcal{V}_T(\Phi)=\mathcal{V}_T(\Phi_0)\\ &\Rightarrow \mathcal{V}_0(\Phi)=\mathcal{V}_0(\Phi_0)\\ &\Rightarrow \mathcal{V}_0(\Phi)=0.\end{aligned}$$

因此 2) 成立. 反过来, 假设任意损益为 0 的交易策略的初始价格为 0, 那么

$$\begin{aligned}\mathcal{V}_T(\Phi_1)=\mathcal{V}_T(\Phi_2) &\Rightarrow \mathcal{V}_T(\Phi_1-\Phi_2)=0\\ &\Rightarrow \mathcal{V}_0(\Phi_1-\Phi_2)=0\\ &\Rightarrow \mathcal{V}_0(\Phi_1)=\mathcal{V}_0(\Phi_2).\end{aligned}$$

因此一价定律成立.

11. 在该情形下, 随机变量 $X_k=1_{\{\omega_k\}}$ 是可达到的, 记其复制策略为 Φ_k. 因此, 任何随机变量 $X=\sum X(\omega_k)X_k$ 都是可达到的, 其复制策略可为 $X=\sum X(\omega_k)\Phi_k$.

13. 其解为

$$\begin{aligned}\theta_{1,1} &= \mathrm{e}^{-rT}\frac{uf_d-df_u}{u-d},\\ \theta_{1,2} &= \frac{f_u-f_d}{S_0(u-d)}.\end{aligned}$$

15. 它是鞅测度的一个选择.

17. 我们有

$$u=\frac{101}{100}=1.01,\quad d=\frac{99}{100}=0.99$$

和

$$f_x=\max(100x-99.50,0)=\begin{cases}1.50, & x=u,\\ 0, & x=d.\end{cases}$$

因此

$$C = \frac{1-(0.99)\mathrm{e}^{-rT}}{0.02}1.50 = 75(1-(0.99)\mathrm{e}^{-rT}) = 0.75813654.$$

19. 一种交易策略 (它不是别的而只是单个的投资组合) 只是二维的, 尽管包含未定权益的空间是三维的. 因此, 估值算子 $\mathcal{V}_1$ 不会是满射的.

练 习 7

1. a) 0, b) 0.0015, c) 0.2944, d) 0.0816, e) 2.0799, f) 0.0783. 对卖权, 利用卖权－买权平价公式 $P = K\mathrm{e}^{-rt} + C - S_0$. 例如, 当 $K = 50$ 时, 我们有

$$P = 50\mathrm{e}^{0.01/6} + 0.2944 - 50 = 0.3778.$$

3. 10% 的收入伴随着 10% 的损失, 反之一样. 结果会有一点轻微的损失, 如下式所示

$$(1+0.1)(1-0.1) = 1 - 0.01 = 0.99.$$

如果收入在先, 则损失的量更大; 如果损失在先, 则收入的量更小.

7. 对 d) 部分, 我们能对路径依赖未定权益 X 定价如下

$$\begin{aligned}\mathcal{I}(X) &= \mathcal{V}_0(\Phi)\\ &= \mathrm{e}^{-rL}\mathcal{E}_\Pi(\mathcal{V}_i(\Phi))\\ &= \mathrm{e}^{-rL}\mathcal{E}_\Pi(X)\\ &= \mathrm{e}^{-rL}\sum_{k=0}^{T} X(\forall\omega\in G_k)\mathbb{P}_\Pi(G_k)\\ &= \mathrm{e}^{-rL}\sum_{k=0}^{T} X_k\mathbb{P}_\Pi(G_k)\end{aligned}$$

练 习 8

5. 如果 $a\in\mathbb{R}$, 那么

$$(-\infty, a) = \bigcup_{n>a}(-n, a).$$

即 $(-\infty, a)$ 为可数开区间的并, 因此它是 Borel 集. 同样

$$(-\infty, a] = \bigcap_{n>0}\Big(-\infty, a+\frac{1}{n}\Big) \in \mathcal{B}.$$

右射线是左射线的补.

7. 可写成

$$A=(A\backslash B)\cup(A\cap B),$$
$$B=(B\backslash A)\cup(A\cap B).$$

因为是不相交的并, 所以

$$\mathbb{P}(A)=\mathbb{P}(A\backslash B)\cup\mathbb{P}(A\cap B),$$
$$\mathbb{P}(B)=\mathbb{P}(B\backslash A)\cup\mathbb{P}(A\cap B),$$

故

$$\mathbb{P}(A)+\mathbb{P}(B)=\mathbb{P}(A\backslash B)+\mathbb{P}(A\cap B)+\mathbb{P}(B\backslash A)+\mathbb{P}(A\cap B).$$

但是

$$(A\backslash B)\cup(A\cap B)\cap(B\backslash A)$$

是不相交的并且等于 $A\cup B$, 从而命题得证.

13. 假设 $A_1\subseteq A_2\subseteq\cdots$ 是递增序列, 令

$$A=\bigcup_{i=1}^{\infty}A_i.$$

该极限存在, 因为它是一个递增有界的实数列的极限. 令 $A_0=\varnothing$, 将 A 写成

$$A=\bigcup_{i=1}^{\infty}(A_i\backslash A_{i-1}),$$

这里事件 $A_i\backslash A_{i-1}$ 互不相交. 那么

$$\begin{aligned}\mathbb{P}(A)&=\mathbb{P}\left(\bigcup_{i=1}^{\infty}(A_i\backslash A_{i-1})\right)\\&=\sum_{i=1}^{\infty}\mathbb{P}(A_i\backslash A_{i-1})\\&=\lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{P}(A_i\backslash A_{i-1})\\&=\lim_{n\to\infty}[\mathbb{P}(A_i)-\mathbb{P}(A_0)]\\&=\lim_{n\to\infty}\mathbb{P}(A_i).\end{aligned}$$

15. 对 b) 部分

$$\begin{aligned}\mathbb{P}((a,b]) &= \mathbb{P}\left(\bigcup_{n=1}^{\infty}\left(a, b-\frac{1}{n}\right]\right)\\ &= \lim_{n\to\infty}\mathbb{P}\left(\left(a, b-\frac{1}{n}\right]\right)\\ &= \lim_{n\to\infty}F\left(b-\frac{1}{n}\right)-F(a)\\ &= F(b-)-F(a).\end{aligned}$$

练 习 9

1. $\mu = 0.15,\ \sigma^2 = 0.03.$
3. $C = \$33.36.$
5. 关于期望, 有

$$\begin{aligned}\mathcal{E}_p(Q_s) &= \mathcal{E}_p\left(\sigma_s\sqrt{\Delta t}\sum_{i=1}^{T}X_{s,i}\right)\\ &= \sigma_s\sqrt{\Delta t}\sum_{i=1}^{T}\mathcal{E}_p(X_{s,i})\\ &= \frac{\sigma_s\sqrt{\Delta t}}{\sqrt{s(1-s)}}\sum_{i=1}^{T}[(1-s)p-s(1-p)]\\ &= \frac{\sigma_s\sqrt{\Delta t}}{\sqrt{s(1-s)}}T(p-s)\\ &= \sigma_s\frac{t}{\sqrt{\Delta t}}\frac{p-s}{\sqrt{s(1-s)}}.\end{aligned}$$

10. 我们有

$$\begin{aligned}S_0 &= \mathrm{e}^{-rt}\mathcal{E}_\Pi(S_t)\\ &= \mathrm{e}^{-rt}\mathcal{E}_\Pi(S_0\mathrm{e}^{\mu_\nu t+\sigma_\nu\sqrt{t}Z_t})\\ &= S_0\mathrm{e}^{-rt+\mu_\nu t}\mathcal{E}_\Pi(\mathrm{e}^{\sigma_\nu\sqrt{t}Z_t})\\ &= S_0\mathrm{e}^{t(\mu_\nu-r)}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{\sigma_\nu\sqrt{t}x}e^{-x^2/2}\mathrm{d}x\\ &= S_0\mathrm{e}^{t(\mu_\nu-r)}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{1}{2}(x^2-2\sigma_\nu\sqrt{t}x+\sigma_\nu^2t)+\frac{1}{2}\sigma_\nu^2t}\mathrm{d}x\\ &= S_0\mathrm{e}^{t(\mu_\nu-r)+\frac{1}{2}\sigma_\nu^2t}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{1}{2}(x-\sigma_\nu\sqrt{t})^2}\mathrm{d}x\\ &= S_0\mathrm{e}^{t(\mu_\nu-r)+\frac{1}{2}\sigma_\nu^2t}.\end{aligned}$$

因此

$$\mathrm{e}^{t(\mu_\nu - r)+\frac{1}{2}\sigma_\nu^2 t} = 1.$$

上式成立当且仅当

$$t(\mu_\nu - r) + \frac{1}{2}\sigma_\nu^2 t = 0.$$

即

$$\mu_\nu = r - \frac{1}{2}\sigma_\nu^2.$$

11. 设 N_U 是一个随机变量, 表示在模型整个周期内股价上涨的次数. 那么, 当然下跌的次数为 $T - N_U$. 从而

$$\begin{aligned} S_{t,T} &= S_0 u^{N_U} d^{T-N_U} \\ &= S_0 \mathrm{e}^{N_U \log u + (T-N_U)\log d} \\ &= S_0 \mathrm{e}^{N_U(\log u - \log d) + T\log d}. \end{aligned}$$

因此

$$H_{t,T} = N_U(\log u - \log d) + T\log d.$$

又因为 N_U 是参数为 T 和 ν 的二项分布, 其期望和方差分别为 $\mathcal{E}(N_U) = T\nu$ 和 $\mathrm{var}(N_U) = T\nu(1-\nu)$. 故

$$\begin{aligned} \mathcal{E}(H_{t,T}) &= T\nu(\log u - \log d) + T\log d, \\ \mathrm{var}(H_{t,T}) &= T\nu(1-\nu)(\log u - \log d)^2. \end{aligned}$$

13. 高波动率暗示着相对于低波动率而言股价更可能偏离执行价. 股价高对买权的持有者来说是好事. 另一方面, 当股价下跌到执行价以下, 而不管下跌了多少 —— 买权将到期, 持有者损失开始购买期权的价钱. 因此, 向上的高波动率是好事, 而向下的高波动率与此无关. 对多头卖权可进行相同的讨论.

练 习 10

3. 首次达到 $S_k \geqslant 2S_0$ 的时刻.

5. 离开一个集合 B 与进入它的补集 B^c 是一样的.

6. 对最大值来说, 我们有

$$[\max\{\tau,\sigma\}=k]=\bigcup_{i=0}^{k}([\tau=i]\cap[\sigma=k-i])\in\mathcal{A}(\mathcal{P}_k)$$

之差不是一个停时, 因为它要求知道未来.

11. 这是一个用 Excel 求解的表.

u=	1.2	r=	0	K=	21	pi=	0.5
d=	0.8	S0=	20	T=	3	Call=1/Put=−1	1

S0	S1	S2	S3	Y0Bar	Y1Bar	Y2Bar	Y3Bar
20	24	28.8	34.56	0	3	7.8	13.56
	16	19.2	23.04		0	0	2.04
		19.2	23.04			0	2.04
		12.8	15.36			0	0
			23.04				2.04
			15.36				0
			15.36				0
			10.24				0

V3Bar	E(V3Bar|P2)	V2Bar	E(V2Bar|P1)	V1Bar	E(V1Bar|P0)	V0Bar
13.56	7.8	7.8	4.41	4.41	2.46	2.46
2.04	7.8	7.8	4.41	4.41	2.46	2.46
2.04	1.02	1.02	4.41	4.41	2.46	2.46
0	1.02	1.02	4.41	4.41	2.46	2.46
2.04	1.02	1.02	0.51	0.51	2.46	2.46
0	1.02	1.02	0.51	0.51	2.46	2.46
0	0	0	0.51	0.51	2.46	2.46
0	0	0	0.51	0.51	2.46	2.46

参考文献

金融数学

[1] Baxter M and Rennie A. Financial Calculus: An Introduction to Derivatives Pricing. Cambridge University Press, 1996, 0-521-55289-3.

[2] Etheridge A. A Course in Financial Calculus. Cambridge University Press, 2002, 0-521-89077-2.

[3] Föllmer H. Stochastic Finance: An Introduction in Discrete Time. Walter de Gruyter, 2002, 3-11-017119-8.

[4] Kallianpur G and Karandikar R. Introduction to Option Pricing Theory. Birkhauser, 1999, 0-8176-4108-4.

[5] Lo A and MacKinlay C. A Non-Random Walk Down Wall Street. Princeton University Press, 2001, 0-691-09256-7.

[6] Medina P and Merino S. Mathematical Finance and Probability. Birkhauser, 2003, 3-7643-6921-3.

[7] Ross S. An Introduction to Mathematical Finance: Options and Other Topics. Cambridge University Press, 1999, 0-521-77043-2.

[8] Ross S. An Elementary Introduction to Mathematical Finance. Cambridge University Press, 2002, 0-521-81429-4.

[9] Shafer G and Vovk V. Probability and Finance: It's Only a Game. John Wiley, 0-471-40226-5.

[10] Wilmott P, Howison S and Dewynne J. The Mathematics of Financial Derivatives: A Student Introduction. Cambridge University Press, 1995, 0-521-49789-2.

概 率 论

[1] Chung K. A Course in Probability Theory. Academic Press, 2001, 0-12-174151-6.

[2] Grimmett Geoffrey. Probability. Oxford University Press, 1986, 0-19-853264-4.

[3] Grimmett Geoffrey R. One Thousand Exercises in Probability. Oxford University Press, 2001, 0-19-857221-2.

[4] Grimmett G and Stirzaker D. Probability and Random Processes. Oxford University Press, 2001, 0-19-857222-0.

[5] Lewis H W. Why Flip a Coin?: The Art and Science of Good Decisions. Wiley, 1997, 0-471-16597-2.

[6] Karr Alan. Probability. Springer, 1993, 0-387-94071-5.

[7] Issac Richard. The Pleasures of Probability. Springer, 1995, 0-387-94415-X.

[8] Shiryaev Albert N. Probability. Springer, 1995, 0-387-94549-0.

[9] Gordon Hugh. Discrete Probability. Springer, 1997, 0-387-98227-2.

[10] Billingsley Patrick. Probability and Measure. Wiley, 1995, 0-471-00710-2.

[11] Feller William. An Introduction to Probability Theory and Its Applications. 3E. Vol. 1. Wiley, 1968, 0-471-25708-7.

[12] Feller William. An Introduction to Probability Theory and Its Applications. 2E. Vol. 2. Wiley, 1971, 0-471-25709-5.

[13] Mises Richard von. Probability, Statistics And Truth. Dover, 0-486-24214-5.

[14] Heathcote C R. Probability: Elements of the Mathematical Theory. Dover, 0-486-41149-4.

[15] Goldberg Samuel. Probability: An Introduction. Dover, 0-486-65252-1.

[16] Mosteller Frederick. Fifty Challenging Problems in Probability With Solutions. Dover, 0-486-65355-2.

[17] Freund John E. Introduction To Probability. Dover, 0-486-67549-1.

[18] Brieman L. Probability. SIAM, 1992, 0-89871-296-3.

[19] Brzezniak Z and Zastawniak T. Basic Stochastic Processes. Springer, 1998, 3-540-76175-6

期权和其他衍生品

[1] Chriss Neil. Black-Scholes and Beyond. McGraw-Hill, 1997, 0-7863-1025-1.

[2] Hull John. Options, Futures and Other Derivatives. Fifth Edition, Prentice Hall, 2003, 0-13-009056-5.

[3] Irwin R. Option Volatility & Pricing. McGraw-Hill, 1994, 1-55738-486-X.

[4] McMillam L. Options as a Strategic Investment. New York Institute of Finance, 2002, 0-7352-0197-8.

[5] Thomsett M. Getting Started in Options. Wiley, 2001, 0-471-40946-4.

附录 A　在不完全市场中对不可达到的未定权益定价

在该附录中，我们讨论在一个不完全离散模型中定价不可达到的未定权益的定价问题. 这些材料不作为该书的主体部分，可在阅读了第 6 章后的任意时间内学习该部分内容.

A.1　不完全市场中的公平价值

经常出现这种情形，一个离散模型是不完全的. 例如，考虑一个单期模型. 在该模型中，一种交易策略就简化为一个投资组合 $\Phi=(\theta_1,\cdots,\theta_n)$，在 t_0 时获得并在模型整个单期内持有. 因此，对每个 i 来说，θ_i 为常数.

现在考虑一个未定权益 $X=(x_1,\cdots,x_m)$. 为了使该未定权益是可达到的，必须存在投资组合 (向量)Φ 使得 $\mathcal{V}_T(\Theta)=X$，也就是说，下面含有 n 个变量 $\theta_1,\cdots,\theta_n$ 的 m 个方程 (m 是状态数) 系统必须满足

$$S_{T,1}(\omega_1)\theta_1+\cdots+S_{T,n}(\omega_1)\theta_n=x_1,$$

$$\cdots\cdots\cdots\cdots$$

$$S_{T,1}(\omega_m)\theta_1+\cdots+S_{T,n}(\omega_m)\theta_n=x_m.$$

如果方程数 m 大于变量数 n，那么不可能对所有可能的向量 $X=(x_1,\cdots,x_m)$ 都存在一个解 (将上述方程系统写成矩阵的形式 $S\Theta=X$，我们看到左边部分定义了一个从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性变换 $f(\Theta)=S\Theta$，所以像 $f(\mathbb{R}^n)$ 的维数至多是 n. 因此，如果 $n<m$，像不可能是整个 $\mathbb{R}^m$ 空间).

故只要 $n<m$，我们就知道市场是不完全的，换句话说，只要经济的状态数大于模型中的资产种类 (在其他时候也一样). 但是一个模型中的状态数多于资产种类也不是一点都不合理.

这就引出了一个问题，如何赋予不完全离散模型中的不可达到的未定权益一个公平价格. 特别地，假设模型 $\mathbb{M}$ 是不完全的，并令 W 是一不可达到的向量，即 $W\notin\mathcal{M}$. 既然不能用复制未定权益的定价程序来赋予 W 一个公平值，那么我们该如何赋予 W 一个公平值呢?

让我们站在一个想出售损益为 W 的投资者这个角度来考虑该问题. 投资者知

道他不可能通过一个自融资策略来复制损益 W, 因为 W 是不可达到的. 但是, 为了对冲空头的风险, 该投资者可以构建一种自融资交易策略 Φ 使它的损益占优 W, 即 $\mathcal{V}_T(\Phi) \geqslant W$. 从此以后, 来自 Φ 的损益就能足够地平仓 W, 可能更多.

另一方面, 考虑所有占优 W 的可达到未定权益组成的集合, 记成

$$\mathcal{D}_W = \{X \in \mathcal{M} | X \geqslant M\}.$$

然后投资者就能通过自融资交易来对冲 W, 其中该策略复制了 $\mathcal{D}_W$ 中的某一向量.

当然, 这里还存在一些问题, 首先 $\mathcal{D}_W$ 是否为空集. 如果 $\mathcal{D}_W$ 是空集, 那么就不存在占优未定权益, 就不能运用上述策略. 先抛开这一点, 假设对冲者在找到一个占优可达到的未定权益 $X \in \mathcal{D}_W$ 上没有困难, 至少从理论上说是这样的. 当然, 所有占优可达到的未定权益 X 有一个公平价格 (无套利价格)$\mathcal{I}(X)$, 这里 $\mathcal{I}$ 为初始定价泛函. 然后我们就能定义 W 的价格为所有占优可达到的未定权益的价格中的最小值, 也就是对冲 W 的风险所需要的最小价值. 用符号表示为

$$P(W) = \min_{X \in \mathcal{D}_W} \mathcal{I}(X).$$

A.2 数学背景

为了在该问题上更深入地展开, 首先得学习一些与线性泛函有关的数学知识.

强正向量形成的基

我们从一个关于强正向量形成的基的结论开始.

定理 1 令 $\mathcal{S}$ 是 $\mathbb{R}^m$ 的一个 k 维子空间, $\mathcal{S}$ 包含一个强正的向量 Z.

1) 所有 "足够接近" Z 的向量也都是强正的. 特别地, 存在一个实数 $\varepsilon > 0$, 使得所有满足 $|X - Z| < \varepsilon$ 的向量 $X \in \mathcal{S}$ 也都是强正的.

2) 存在一组强正向量是 $\mathcal{S}$ 的基.

证明 1) 记 $Z = (z_1, \cdots, z_m)$, 并令 $\varepsilon = \min\{z_i\}$. 如果 $|X - Z| < \varepsilon$, 那么

$$z_i - x_i \leqslant |z_i - x_i| \leqslant |Z - X| < \varepsilon \leqslant z_i.$$

因此对所有的 i, 必须有 $x_i > 0$, 故 X 是强正的.

2) 想法主要是, Z 的一个充分小的球包含了足够多的 "方向" 来定义一个基, 再根据定理的第一部分知这些向量都是强正的. 特别地, 记 $Z = Z_1$, 将 $\{Z_1\}$ 扩展成 $\mathcal{S}$ 的基 $\mathcal{B} = \{Z_1, \cdots, Z_k\}$. 考虑向量

$$Z_i(\lambda) = \lambda Z_1 + (1 - \lambda) Z_i,$$

对 $i=1,\cdots,k$ 和 $\lambda\in\mathbb{R}$. 我们断定, 对任意 $\lambda\in(0,1)$ 这些向量线性独立, 因此形成了 $\mathcal{S}$ 的一个基. 为了看清这点, 假设

$$\sum_{i=1}^{k} a_i Z_i(\lambda)=0,$$

其中 a_i 为实数点. 那么

$$\begin{aligned}0&=\sum_{i=1}^{k} a_i Z_i(\lambda)\\&=\sum_{i=1}^{k} a_i(\lambda Z_1+(1-\lambda)Z_i)\\&=\left[\sum_{i=1}^{k} a_i\lambda\right] Z_1+\sum_{i=1}^{k} a_i(1-\lambda)Z_i.\end{aligned}$$

又由于 $\mathcal{B}$ 是一个基, 故对所有的 $i\geqslant 2$ 有 $a_i(1-\lambda)=0$, 这就表明对所有的 i, 有 $a_i=0$.

最后, 注意到

$$|Z_i(\lambda)-Z_1|=|\lambda Z_1+(1-\lambda)Z_i-Z_1|=(1-\lambda)|Z_i-Z_1|.$$

因此可选 λ 充分接近 1, 再应用该定理第一部分结论就能得出向量 $Z_i(\lambda)$ 是强正的. 定理得证.

占优向量

我们现在将精力放到占优向量什么时候存在这个问题上来.

定理 2 如果子空间 $\mathcal{S}\subseteq\mathbb{R}^m$ 包含一个强正向量, 那么对任何 $Z\in\mathbb{R}^m$, 集合 $\mathcal{D}_Z=\{X\in\mathcal{S}|X\geqslant Z\}$ 非空.

证明 设 $\mathcal{S}$ 包含一个强正向量 $A=(a_1,\cdots,a_m)$, 即对所有的 i 有 $a_i>0$. 然后我们需要做的就只是将 A 乘以一个充分大的 α, 使得对每个 i, 有 $\alpha a_i\geqslant z_i$. 我们将细节留给读者.

可达到的向量空间

下面的结论指出了我们为什么需要先陈述上面两个定理.

定理 3 离散模型中所有可达到向量形成的集合 $\mathcal{M}$ 包含一个强正向量. 因此

1) $\mathcal{M}$ 有一个由强正向量形成的基.

2) 对任何 $Z\in\mathbb{R}^m$, 集合 $\mathcal{D}_Z=\{X\in\mathcal{M}|X\geqslant Z\}$ 非空.

证明 令 Φ 为这样的一种交易策略, 在无风险资产上投资一个单位, 然后一直持有下去. 那么最终的损益 $X = \mathcal{V}_T(\Phi) \in \mathcal{M}$ 是强正的.

线性泛函

为了提供查询, 这里给出线性泛函的定义, 可能大家都非常熟悉.

定义 1 令 $\mathcal{S}$ 是 $\mathbb{R}^m$ 的一个子空间. $\mathcal{S}$ 上的一个线性泛函是这样的一个函数 $f: \mathcal{S} \to \mathbb{R}$, 对所有的 $a, b \in \mathbb{R}$ 和 $X, Y \in \mathcal{S}$ 满足

$$f(aX + bY) = af(X) + bf(Y).$$

既然一个线性泛函的像仅是一维的, 故它的核"非常大", 即

$$\dim(\ker(f)) = \dim(\mathcal{S}) - 1.$$

因此, 存在一个非零向量 W, 使得

$$\mathcal{S} = \ker(f) \oplus \langle W \rangle,$$

这里直和是正交的, 即 $W \perp \ker(f)$. 因此, 任何向量 $X \in \mathbb{R}^m$ 都具有 $X = Z + aW$ 这种形式, 这里 $Z \in \ker(f), a \in \mathbb{R}$. 从而

$$f(X) = f(Z + aW) = f(Z) + af(W) = af(W).$$

这表明 f 完全由它在向量 W 上的值决定.

有一些关于上面结论的有用推论. 例如, 如果 g 是 $\mathcal{S}$ 上的另一线性泛函且 g 与 f 有相同的核, 从而也有 $g(X) = ag(W)$ 且 (假设 f 非零)

$$g(X) = ag(W) = a\frac{g(W)}{f(W)}f(W) = \frac{g(W)}{f(W)}f(X) = \lambda f(X),$$

这里 $\lambda = g(W)/f(W)$ 为常数. 我们已经证明了下面的定理.

定理 4 令 f 和 g 是子空间 $\mathcal{S} \subseteq \mathbb{R}^m$ 上的线性泛函. 如果 $\ker(f) = \ker(g)$, 那么存在一个实数 λ, 使得 $g = \lambda f$.

线性泛函的表示定理

下面这个非常著名的定理利用内积从单个向量方面刻画了线性泛函. 我们记 $\mathbb{R}^m$ 的一个标准基为 $E_1, \cdots, E_m$.

定理 5 (线性泛函的表示定理) 1) 令 $f: \mathcal{S} \to \mathbb{R}$ 为子空间 $\mathcal{S} \subseteq \mathbb{R}^m$ 上的线性泛函. 那么存在唯一的向量 $Y_f \in \mathcal{S}$, 使得对所有的 $X \in \mathcal{S}$, 有

$$f(X) = \langle X, Y_f \rangle.$$

2) 如果 $f: \mathbb{R}^m \to \mathbb{R}$ 是 $\mathbb{R}^m$ 上的线性泛函, 那么

$$Y_f = \langle f(E_1), \cdots, f(E_m) \rangle.$$

证明 1) 设 $W \in (\ker(f))^{\perp}$ 的长度为单位 1, 且令 $Y_f = f(W)W$. 因为任何 $X \in \mathcal{S}$ 都具有形式 $X = Z + \alpha W$, 其中 $Z \in \ker(f)$, 较容易地知道 $f(Z) = \langle Z, Y_f \rangle$ (因为两边都为 0) 和 $f(\alpha W) = \langle \alpha W, Y_f \rangle$ (因为 $\langle W, W \rangle = 1$), 所以对所有的 $Z \in \mathcal{S}$ 有 $f(Z) = \langle Z, Y_f \rangle$, 这就是所要的结果. 为了证明唯一性, 假设对所有的 $X \in \mathcal{S}$, 有

$$f(X) = \langle X, Y \rangle = \langle X, Z \rangle.$$

那么对所有的 $X \in \mathcal{S}$, 有

$$\langle X, Y - Z \rangle = 0.$$

取 $X = Y - Z$ 得到

$$\langle Y - Z, Y - Z \rangle = 0.$$

这仅当 $Y - Z = 0$, 即 $Y = Z$ 时上式才成立. 我们将 2) 的证明留给读者.

线性泛函的延拓

如果线性泛函 f 定义在 $\mathbb{R}^m$ 上的一个合适子空间 $\mathcal{S}$ 上, 那么有无穷多种方法可以将 f 扩展为整个 $\mathbb{R}^m$ 上的线性泛函. 事实上, 如果 $X_1, \cdots, X_k$ 是 $\mathcal{S}$ 上的一个基, 那么我们可以把该基扩展为 $\mathbb{R}^m$ 上的基 $X_1, \cdots, X_k, Y_1, \cdots, Y_{m-k}$, 然后就能在向量 Y_i 上以任何一种方式定义 f. 结果就唯一确定了 $\mathbb{R}^m$ 上一个独特的线性泛函.

我们想更深入地探究延拓这个概念, 但是首先需给出延拓的一个正式定义.

定义 2 设 $f: \mathcal{S} \to \mathbb{R}$ 为子空间 $\mathcal{S} \subseteq \mathbb{R}^m$ 上的线性泛函. 那么 f 关于 $\mathbb{R}^m$ 的延拓是一个线性泛函 $\overline{f}: \mathbb{R}^m \to \mathbb{R}$, 对所有的 $X \in S$ 满足

$$\overline{f}(X) = f(X).$$

下一定理刻画了线性泛函的两种延拓方法.

定理 6 设 $f: \mathcal{S} \to \mathbb{R}$ 为子空间 $\mathcal{S} \subseteq \mathbb{R}^m$ 上的线性泛函, 且 $K = \ker(f) \subseteq \mathcal{S}$.

1) 那么 $g: \mathbb{R}^m \to \mathbb{R}$ 是 f 的延拓当且仅当对某个 $W \in \mathcal{S}^{\perp}$, 有

$$Y_g = Y_f + W,$$

即

$$g(X) = \langle X, Y_f + W \rangle,$$

这里 $W \in \mathcal{S}^\perp$.

2) 如果 f 非零且 $g : \mathbb{R}^m \to \mathbb{R}$ 是 f 的延拓, 那么 $Y_g \in K^\perp \backslash \mathcal{S}^\perp$. 而且, 对每个向量 $U \in K^\perp \backslash \mathcal{S}^\perp$, 存在唯一的标量 λ, 使得线性泛函 $g(X) = \langle X, \lambda U \rangle$ 是 f 的延拓.

证明 1) 假设 $g(X) = \langle X, Y_g \rangle$. 那么说 g 是 f 的一个延拓等价于说对所有的 $S \in \mathcal{S}$, 有

$$\langle S, Y_g \rangle = \langle S, Y_f \rangle,$$

或者

$$\langle S, Y_g - Y_f \rangle = 0.$$

换句话说, $Y_g - Y_f \in \mathcal{S}^\perp$ 或等价 $Y_g = Y_f + W$, 其中 $W \in \mathcal{S}^\perp$.

2) 根据 1) 可得出第一部分, 因为 $Y_g = Y_f + W$, 其中 $W \in \mathcal{S}^\perp$. 现在 $Y_f \in K^\perp$ 和 $W \in \mathcal{S}^\perp \subseteq K^\perp$, 因而 $Y_g \in K^\perp$. 另一方面, 如果 $Y_g \in \mathcal{S}^\perp$, 那么也有 $Y_f \in \mathcal{S}^\perp$, 这就表明 $f = 0$, 与假设矛盾. 故 $Y_g \in K^\perp \backslash \mathcal{S}^\perp$.

对于第二部分, 令 $U \in K^\perp \backslash \mathcal{S}^\perp$, 定义 $g(X) = \langle X, \lambda U \rangle$. 记 g 在 $\mathcal{S}$ 上的限制为 $g|_\mathcal{S}$. 首先指出有

$$\ker(g|_\mathcal{S}) = \ker(f).$$

首先, 注意到

$$\ker(g|_\mathcal{S}) = \ker(g) \cap \mathcal{S} = \langle U \rangle^\perp \cap \mathcal{S}.$$

因为 $U \in K^\perp$, 所以有 $\langle U \rangle \subseteq K^\perp$ 和 $K \subseteq \langle U \rangle^\perp$. 同样因为 $K \subseteq \mathcal{S}$, 故 $K \subseteq \langle U \rangle^\perp \cap \mathcal{S}$. 为了完成等式的证明, 我们指出这两个子空间的维数相同. 由于 $U \notin \mathcal{S}^\perp$, 从而知道 $\mathcal{S} \nsubseteq \langle U \rangle^\perp$, 故

$$\dim(\langle U \rangle^\perp + \mathcal{S}) > \dim(\langle U \rangle^\perp) = m - 1.$$

这表明 $\dim(\langle U \rangle^\perp + \mathcal{S}) = m$. 因此, 根据线性代数里一个著名的公式, 有

$$\begin{aligned}
\dim(\langle U \rangle^\perp \cap \mathcal{S}) &= \dim(\langle U \rangle^\perp) + \dim(\mathcal{S}) - \dim(\langle U \rangle^\perp + \mathcal{S}) \\
&= m - 1 + \dim(\mathcal{S}) - m \\
&= \dim(\mathcal{S}) - 1 \\
&= \dim(\ker(f)).
\end{aligned}$$

因而 $K = \langle U \rangle^\perp \cap \mathcal{S}$, 即有 $\ker(g|_\mathcal{S}) = \ker(f)$. 因此由定理 4 知存在一个标量 μ, 使得 $g|_\mathcal{S} = \mu f$. 如果 $\lambda = \dfrac{1}{\mu}$, 那么有 $f = (\lambda g)|_\mathcal{S}$, 也就是说 λg 是 f 的一个延拓.

为了完成证明, 我们只需指出这样的 λ 是唯一的. 假设 λ_1 和 λ_2 具有性质: $g_1(X) = \langle X, \lambda_1 U\rangle$ 和 $g_2(X) = \langle X, \lambda_2 U\rangle$ 都是 f 的延拓. 由于对所有的 $v \in \mathcal{S}$, 有 $g_1(X) = g_2(X)$, 从而对所有的 $X \in \mathcal{S}$, 有 $\langle X, (\lambda_1 - \lambda_2)U\rangle = 0$. 这意味着 $(\lambda_1 - \lambda_2)U \in \mathcal{S}^{\perp}$. 但是又因为 $U \notin \mathcal{S}^{\perp}$, 所以推出 $\lambda_1 - \lambda_2 = 0$, 即 $\lambda_1 = \lambda_2$, 得证.

正线性泛函

我们已经定义了向量的正性. 对于线性泛函, 有下面的定义.

定义 3　设 $f: \mathcal{S} \to \mathbb{R}$ 为子空间 $\mathcal{S} \subseteq \mathbb{R}^m$ 上的线性泛函. 那么

1) f 是非负的 (记$f \geqslant 0$), 如果对所有的 $X \in \mathcal{S}$, 有

$$X \geqslant 0 \Rightarrow f(X) \geqslant 0.$$

2) f 是严格正的 (记$f > 0$), 如果对所有的 $X \in \mathcal{S}$, 有

$$X > 0 \Rightarrow f(X) > 0.$$

不难发现初始定价泛函 $\mathcal{I}$ 是严格正的 (在无套利下). 如果 $X > 0$ 是一个可达到的未定权益且交易策略 Φ 复制了 X, 那么 $\mathcal{V}_T(\Phi) = X > 0$, 因此初始价格

$$\mathcal{I}(X) = \mathcal{V}_0(X)$$

必须是正的, 否则就存在套利机会. 这就是我们为什么要学习严格正的线性泛函的原因.

下一定理刻画了 $\mathbb{R}^m$ 上线性泛函的正性.

定理 7　设 $f: \mathbb{R}^m \to \mathbb{R}$ 是一个非零线性泛函. 那么

1) f 是非负的当且仅当 Y_f 是严格正的, 即

$$Y_f = \langle f(E_1), \cdots, f(E_m)\rangle > 0.$$

2) f 是严格正的当且仅当 Y_f 是强正的, 即

$$Y_f = \langle f(E_1), \cdots, f(E_m)\rangle \gg 0.$$

证明　1) 首先假设 f 是非负的. 由 $E_i \geqslant 0$ 得 $f(E_i) \geqslant 0$. 因为 $f(E_i)$ 不可能全为 0, 否则 f 就为零, 故推出 $Y_f > 0$. 反之, 如果对所有的 i, 有 $f(E_i) \geqslant 0$(不全为 0), 又因为任何 $X \geqslant 0$ 可写成

$$X = (x_1, \cdots, x_m),$$

其中 $x_i \geqslant 0$, 故可推得

$$f(X) = \sum_{i=1}^{m} x_i f(E_i) \geqslant 0.$$

2) 的证明相似, 我们将证明留给读者.

值得指出的是, 该定理仅适用于全空间 $\mathbb{R}^m$ 上的线性泛函. 例如, 考虑 $\mathbb{R}^2$ 中的一个子空间 $\mathcal{S} = \{(a,-a)|a \in \mathbb{R}\}$. 那么对任意非零向量 $W \in \mathbb{R}^2$, 根据 $g(X) = \langle X, W\rangle$ 定义的线性泛函 $g: \mathcal{S} \to \mathbb{R}$ 是非负的, 因为零向量是 $\mathcal{S}$ 中唯一满足 $X \geqslant 0$ 的向量. 事实上, f 是严格正的, 这是因为 $\mathcal{S}$ 中没有向量满足 $X > 0$!

延拓和正性

我们考虑这样一个问题, 一个子空间 $\mathcal{S}$ 上的严格正的线性泛函能否延拓为 $\mathbb{R}^m$ 上的一个严格正的线性泛函. 对任何线性泛函 f, 令

$$
\begin{aligned}
\mathcal{E}_{\geqslant 0}(f) &= \{f\text{的非负延拓}\},\\
\mathcal{E}_{>0}(f) &= \{f\text{的严格正延拓}\}.
\end{aligned}
$$

因为严格正的线性泛函也是非负的, 故有

$$\mathcal{E}_{>0}(f) \subseteq \mathcal{E}_{\geqslant 0}(f).$$

定理 8 设 $f: \mathcal{S} \to \mathbb{R}$ 为 $\mathcal{S} \subseteq \mathbb{R}^m$ 上严格正的线性泛函. 那么 f 有一个到 $\mathbb{R}^m$ 上严格正的延拓, 即集合 $\mathcal{E}_{>0}(f)$ 非空, 因此 $\mathcal{E}_{\geqslant 0}(f)$ 也非空.

证明 因为对所有的 $X \in \mathcal{S}$, 有

$$X > 0 \Rightarrow f(X) > 0.$$

所以如果 $f(Z) = 0$, 则 $Z \not> 0$. 换句话说

$$\ker(f) \cap \mathbb{R}^m_+ = \{0\},$$

其中

$$\mathbb{R}^m_+ = \{(x_1, \cdots, x_m)|x_i \geqslant 0\}$$

为 $\mathbb{R}^m$ 的非负部分. 由附录 B 中的定理 5 可知 $(\ker(f))^\perp$ 包含一个强正的向量 $W \gg 0$. 然后由定理 1 知, $(\ker(f))^\perp$ 有一个由强正向量组成的基 $\mathcal{B} = \{W_1, \cdots, W_k\}$.

我们声称至少有一个强正的向量 $W_i \in \mathcal{B}$ 不属于 $\mathcal{S}^\perp$. 因为 $\dim(\ker(f)) = \dim(\mathcal{S}) - 1$, 故有

$$
\begin{aligned}
\dim(\mathcal{S}) + \dim(\mathcal{S}^\perp) &= m\\
&= \dim(\ker(f)) + \dim(\ker(f)^\perp)\\
&= \dim(\mathcal{S}) - 1 + \dim(\ker(f)^\perp).
\end{aligned}
$$

从而

$$\dim(\ker(f)^{\perp}) = \dim(\mathcal{S}^{\perp}) + 1.$$

这表明 $(\ker(f)^{\perp})$ 的整个基不可能都在 $\mathcal{S}^{\perp}$ 里. 因此, 对某个 i, 有 $W_i \in (\ker(f))^{\perp} \backslash \mathcal{S}^{\perp}$. 现由定理 6 知, 存在一个标量 λ, 满足

$$g(X) = \langle X, \lambda W_i \rangle,$$

为 f 的延拓. 由于 $f \gg 0$ 和 $W \gg 0$, 则有

$$\lambda \langle W_i, W_i \rangle = \langle W_i, \lambda W_i \rangle = g(W_i) = f(W_i) > 0.$$

这表明 $\lambda > 0$. 因此, λW_i 是强正的, 故根据定理 7 知, $\mathbb{R}^m$ 上的线性泛函 g 是 f 的严格正的延拓.

A.3　对不可达到的未定权益定价

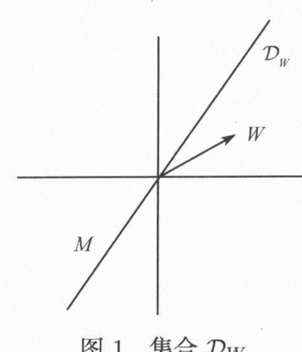


图 1　集合 $\mathcal{D}_W$

现在已经具备了必要的背景知识, 让我们回到对不可达到的未定权益定价这个问题上来. 图 1 给出了我们想要考虑的情形 (至少在两状态情形下). 我们想要定价不可达到的未定权益 W. 占优 W 的可达到未定权益由粗线标出.

回顾一下, 所有占优 W 的可达到未定权益可记成

$$\mathcal{D}_W = \{X \in \mathcal{M} | X \geqslant W\}.$$

然后一个投资者可采用复制了 $\mathcal{D}_W$ 中的任一向量的所有自融资交易策略中的任意一种来对冲 W.

既然可用初始定价泛函 $\mathcal{I}$ 来轻易对 $\mathcal{D}_W$ 中的未定权益进行定价, 那么将 W 的公平价格定为所有占优可达到的未定权益的价格的最小值看起来是合理的, 即对冲 W 的风险所要求的最小价格.

定义 4　对一个不可达到的未定权益 W, 定义最小占优价格为

$$P(W) = \min_{X \in \mathcal{D}_W} \mathcal{I}(X).$$

现在, 因为由可达到的未定权益形成的占优集是一个无限集, 所以我们不知道是否存在一个特殊的占优未定权益 $X \in \mathcal{D}_W$ 确实达到了 $P(W)$ (可类比由所有大于零的实数形成的集合, 在该集合中不存在实数达到最小值零).

因此, 我们遇到了一个问题, 即是否存在 $X_0 \in \mathcal{D}_W$, 使得

$$\mathcal{I}(X_0) = P(W) = \min_{X \in \mathcal{D}_W} \mathcal{I}(X).$$

在继续考虑这个问题之前, 我们稍微换一种观点来思考. 初始定价泛函 $\mathcal{I}$ 仅仅适用于 $\mathcal{M}$ 中的向量, 但是的确可能将 $\mathcal{I}$ 延拓为整个 $\mathbb{R}^m$ 上的线性泛函 $\overline{\mathcal{I}}$, 这可能有无穷多种方法. 将 W 的价格定为其中一个延拓的 $\overline{\mathcal{I}}(W)$ 会怎么样呢?

由于初始价格泛函 $\mathcal{I}$ 是严格正的, 或许我们应该将注意力只放在 $\mathcal{I}$ 的严格正的延拓上.

定义 5　对一个不可达到的未定权益 W, 定义其最大延拓价格为

$$P'(W) = \max_{\overline{\mathcal{I}} \in \mathcal{E}_{>0}(\mathcal{I})} \{\overline{\mathcal{I}}(W)\}.$$

关于这种定价策略, 我们也可以问是否存在 $\mathcal{I}$ 的一种严格正的延拓 $\overline{\mathcal{I}}$ 达到上述最大值, 即

$$\overline{\mathcal{I}}_0(W) = \max_{\overline{\mathcal{I}} \in \mathcal{E}_{>0}(\mathcal{I})} \{\overline{\mathcal{I}}(W)\}.$$

在该情形下, 答案是否定的, 下面的例子说明了这点.

例 7　设 $\mathcal{M} = \{(x,x) | x \in \mathbb{R}^2\} \subseteq \mathbb{R}^2$, 并令 $\mathcal{I}((x,x)) = ax$, 其中 $a > 0$. 那么 $\mathcal{I}((x,x)) = \left\langle (x,x), \left(\frac{a}{2}, \frac{a}{2}\right) \right\rangle$, 因此 $Y_{\mathcal{I}} = \left(\frac{a}{2}, \frac{a}{2}\right)$. $\mathcal{I}$ 的任意延拓 $\overline{\mathcal{I}}$ 具有形式

$$\overline{\mathcal{I}}(X) = \langle X, Y_{\mathcal{I}} + W \rangle,$$

其中 $W \in \mathcal{M}^{\perp}$. 因此, W 具有形式 $\left(\frac{z}{2}, -\frac{z}{2}\right)$, $z \in \mathbb{R}$. 有

$$Y_{\overline{\mathcal{I}}} = \left(\frac{a}{2}, \frac{a}{2}\right) + \left(\frac{z}{2}, -\frac{z}{2}\right) = \frac{1}{2}(a+z, a-z).$$

故 $\overline{\mathcal{I}}$ 非负当且仅当 $-a \leqslant z \leqslant a$, $\overline{\mathcal{I}}$ 严格正当且仅当 $-a < z < a$. 现在令 $W = (0,1)$, 这是不可达到的. 那么

$$\overline{\mathcal{I}}(W) = \langle W, Y_{\overline{\mathcal{I}}} \rangle = \frac{1}{2}(a-z).$$

所以

$$P'(W) = \max_{\overline{\mathcal{I}} \in \mathcal{E}_{\geqslant 0}(\mathcal{I})} \{\overline{\mathcal{I}}(W)\} = \max_{-a \leqslant z \leqslant a} \left\{\frac{1}{2}(a-z)\right\} = a.$$

但是没有严格正的延拓能达到该最大值. 事实上, 能达到该最大值的唯一延拓仅仅是非负的, 此时有 $z = -a$.

考虑到前面这个例子, 我们想放宽自己的视野而包括非负的延拓. 如已经看到的那样, 我们得到相同的最大价格, 即

$$\max_{\overline{\mathcal{I}}\in\mathcal{E}_{>0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}=\max_{\overline{\mathcal{I}}\in\mathcal{E}_{\geqslant 0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}.$$

更进一步, 最大值总是由非负延拓达到. 但是, 我们在这里就不证明这些结论了.

定价问题的最优解

因为, 对于一个不可达到的未定权益 W, 我们有两种定价方法：最小占优价格

$$P(W)=\min_{X\in\mathcal{D}_W}\mathcal{I}(X)$$

和最大延拓价格

$$P'(W)=\max_{\overline{\mathcal{I}}\in\mathcal{E}_{\geqslant 0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}.$$

我们说上述两种价格都是可以达到的. 而且, 令人满意的是上述两种价格相等

$$P(W)=P'(W).$$

下一定理总结了我们刚才讨论的内容.

定理 9　*设 W 是一个不可达到的未定权益.*

1) 集合 $\mathcal{E}_{>0}(\mathcal{I}),\mathcal{E}_{\geqslant 0}(\mathcal{I})$ 和 $\mathcal{D}_W$ 是非空的, 且

$$\max_{\overline{\mathcal{I}}\in\mathcal{E}_{>0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}=\max_{\overline{\mathcal{I}}\in\mathcal{E}_{\geqslant 0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}=\min_{X\in\mathcal{D}_W}\mathcal{I}(X).$$

2) 存在一个可达到的占优未定权益 $X_0\in\mathcal{M}$, 满足

$$\mathcal{I}(X_0)=\min_{X\in\mathcal{D}_W}\mathcal{I}(X).$$

3) 存在一个 $\mathcal{I}$ 的非负延拓 $\overline{\mathcal{I}}$, 满足

$$\overline{\mathcal{I}}_0(W)=\max_{\overline{\mathcal{I}}\in\mathcal{E}_{\geqslant 0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}.$$

所以, 不管是用可到达的占优未定权益还是用初始定价泛函 $\mathcal{I}$ 的延拓来定价不可达到的未定权益, 结果都是一样的. 因此, 我们可以作出如下的定义.

定义 6　*定义一个不可达到的未定权益 W 的公平价格为*

$$\max_{\overline{\mathcal{I}}\in\mathcal{E}_{\geqslant 0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}=\min_{X\in\mathcal{D}_W}\mathcal{I}(X).$$

定理 9 的证明有一点复杂, 需要用到线性规划的知识. 既然我们假定部分读者不熟悉线性规划方面的内容, 并且因为必要的背景知识会将我们带入一个更远的领域, 所以就不证明定理 9 了. 在练习中, 我们要求读者证明

$$\max_{\overline{\mathcal{I}}\in\mathcal{E}_{>0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}=\max_{\overline{\mathcal{I}}\in\mathcal{E}_{\geqslant 0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}$$

和

$$\max_{\overline{\mathcal{I}}\in\mathcal{E}_{\geqslant 0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}\leqslant\min_{X\in\mathcal{D}_W}\mathcal{I}(X).$$

练　习

1. 假设 f 是 $\mathbb{R}^2$ 上的一个线性泛函, 满足：对所有的 $x\in\mathbb{R}^2$,

$$x\gg 0\Rightarrow f(x)>0.$$

证明 f 不一定是严格正的.

2. 如果 $f:\mathbb{R}^m\to\mathbb{R}$ 是 $\mathbb{R}^m$ 上的线性泛函, 那么

$$Y_f=\langle f(E_1),\cdots,f(E_m)\rangle,$$

其中 E_i 是 $\mathbb{R}^m$ 的标准基.

3. 在例 1 中, 证明

$$\max_{\overline{\mathcal{I}}\in\mathcal{E}_{\geqslant 0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}=\min_{X\in\mathcal{D}_W}\mathcal{I}(X).$$

4. 令 $\mathcal{M}=\{(2a,3a)|a\in\mathbb{R}\}\subseteq\mathbb{R}^2$ 和 $\mathcal{I}(2,3)=\alpha$. 证明任意未定权益 $W=(x,y)$ 的价格为

$$\max\left\{\frac{\alpha x}{2},\frac{\alpha y}{3}\right\}.$$

5. 设 W 是一个不可达到的未定权益. 证明

$$\max_{\overline{\mathcal{I}}\in\mathcal{E}_{\geqslant 0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}\leqslant\min_{X\in\mathcal{D}_W}\mathcal{I}(X).$$

提示：首先证明如果有 $X\in\mathcal{D}_W$ 和 $\overline{\mathcal{I}}\in\mathcal{E}_{\geqslant 0}(\mathcal{I})$, 那么

$$\overline{\mathcal{I}}(W)\leqslant\mathcal{I}(X).$$

6. 证明：

$$\max_{\overline{\mathcal{I}}\in\mathcal{E}_{>0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}=\max_{\overline{\mathcal{I}}\in\mathcal{E}_{\geqslant 0}(\mathcal{I})}\{\overline{\mathcal{I}}(W)\}.$$

提示：记 $\mathcal{E}_{\geqslant 0}(\mathcal{I})$ 上的最大值为 M. 那么存在 $\mathcal{E}_{\geqslant 0}(\mathcal{I})$ 中的延拓序列 $\overline{\mathcal{I}}_n$ 满足

$$\lim_{n\to\infty}\overline{\mathcal{I}}_n(W)=M.$$

如果 $g\in\mathcal{E}_{>0}(\mathcal{I})$, 那么可证明对任意 $n>0$, 线性泛函

$$f_n=\Big(1-\frac{1}{n}\Big)\overline{\mathcal{I}}_n+\frac{1}{n}g$$

是 $\mathcal{I}$ 的严格正的延拓. 当 $n\to\infty$ 时 $f_n(W)$ 的极限是什么? 如何证明该结果?

附录 B　凸性和分离定理

在该附录中, 我们将学习凸性方面的一些必要知识.

B.1　凸集, 闭集和紧集

我们需要下面的一些概念.

定义 1　1) 设 $x_1, \cdots, x_k \in \mathbb{R}^n$. 称任何线性组合

$$t_1x_1 + \cdots + t_kx_k,$$

其中

$$t_1 + \cdots + t_k = 1, \quad 0 \leqslant t_i \leqslant 1$$

为向量 $x_1, \cdots, x_k$ 的凸组合.

2) 称子集 $X \subseteq \mathbb{R}^n$ 是凸的, 如果只要 $x, y \in X$, 就有 x 和 y 之间的整个线段部分也在 X 里面, 用符号表示为

$$\{sx + ty |\ s + t = 1, 0 \leqslant s, t \leqslant 1\} \subseteq X.$$

3) 称子集 $X \subseteq \mathbb{R}^n$ 是一个锥, 如果 $x \in X$ 暗示着对所有的 $\alpha > 0$, 有 $\alpha x \in X$.

4) 称子集 $X \subseteq \mathbb{R}^n$ 是闭的, 如果只要 $x_n \in X$ 是 X 中的收敛序列, 就有极限点也在 X 里. 简单地说, 一个子集是闭的如果它对取极限是封闭的.

5) 称子集 $X \subseteq \mathbb{R}^n$ 是紧的, 如果它既是闭的又是有界的.

我们也需要下列的分析结果:

1) 定义在紧集 $X \subseteq \mathbb{R}^n$ 上的连续函数在 X 上可取到最大值和最小值.

2) 子集 $X \subseteq \mathbb{R}^n$ 是紧的当且仅当 X 中的每个序列都有一个子列在 X 中收敛.

定理 1　设 X 和 Y 是 $\mathbb{R}^n$ 中的子集. 定义

$$X + Y = \{a + b |\ a \in X, b \in Y\}.$$

1) 如果 X 和 Y 是凸的, 那么 $X + Y$ 也是凸的.

2) 如果 X 是紧的, Y 是闭的, 那么 $X + Y$ 是闭的.

证明　1) 设 $a_0 + b_0$ 和 $a_1 + b_1$ 是 $X + Y$ 中的元素. 该两点间的线段为

$$t(a_0 + b_0) + (1 - t)(a_1 + b_1) = ta_0 + (1 - t)a_1 + tb_0 + (1 - t)b_1 \in X + Y.$$

因此 $X+Y$ 是凸的.

2) 设 a_n+b_n 是 $X+Y$ 中的收敛序列. 假设 $a_n+b_n\to c$. 我们必须证明 $c\in X+Y$. 因为 a_n 是紧集 X 中的一个序列, 故存在收敛子列 a_{n_k}, 它的极限 α 在 X 里. 由于有 $a_{n_k}+b_{n_k}\to c$ 和 $a_{n_k}\to\alpha$, 故可推出 $b_{n_k}\to c-\alpha$. 又由于 Y 是闭的, 故一定会有 $c-\alpha\in Y$. 因此 $c=\alpha+(c-\alpha)\in X+Y$, 得证.

B.2 凸　　包

我们将需要用到凸包的概念.

定义 2 集合 $S=\{x_1,\cdots,x_k\}\subseteq\mathbb{R}^n$ 的凸包指的是 $\mathbb{R}^n$ 中包含向量 $x_1,\cdots,x_k$ 的最小凸集. 记 S 的凸包为 $\mathcal{C}(S)$.

下面给出凸包的一个刻画.

定理 2 设 $S=\{x_1,\cdots,x_k\}$ 是 $\mathbb{R}^n$ 中的一个集合. 那么凸包 $\mathcal{C}(S)$ 是由 S 的所有凸组合形成的集合

$$\Delta=\left\{t_1x_1+\cdots+t_kx_k \mid 0\leqslant t_i\leqslant 1,\sum t_i=1\right\}.$$

证明 首先, 我们证明 Δ 是凸集. 令

$$X=t_1x_1+\cdots+t_kx_k,$$
$$Y=s_1x_1+\cdots+s_kx_k$$

是 S 的凸组合, 且 $a+b=1, 0\leqslant a,b\leqslant 1$. 那么

$$\begin{aligned}aX+bY&=a(t_1x_1+\cdots+t_kx_k)+b(s_1x_1+\cdots+s_kx_k)\\&=(at_1+bs_1)x_1+\cdots+(at_k+bs_k)x_k.\end{aligned}$$

但是这也是 S 的一个凸组合, 因为有

$$0\leqslant at_i+bs_i\leqslant\begin{cases}(a+b)s_i=s_i\leqslant 1,\ t_i\leqslant s_i,\\(a+b)t_i=t_i\leqslant 1,\ t_i>s_i\end{cases}$$

和

$$\sum_{i=1}^k(at_i+bs_i)=a\sum_{i=1}^k t_i+b\sum_{i=1}^k s_i=a+b=1.$$

因此

$$X,Y\in\Delta\Rightarrow aX+bY\in\Delta.$$

这表明 Δ 是凸集. 另外非常明显的是所有的 $x_i \in \Delta$, 因此 Δ 是凸集且包含了 S 中的所有向量. 从而

$$\mathcal{C}(S) \subseteq \Delta.$$

为了证明反向包含关系, 必须指出任何包含 S 的凸集也一定包含 Δ. 故假设 D 是包含 S 的凸集. 因而 D 包含 S 中任意两个向量的所有凸组合. 考虑一个凸组合, 比方说由三个向量形成的 $t_1x_1 + t_2x_2 + t_3x_3$. 可以写成

$$t_1x_1 + t_2x_2 + t_3x_3 = t_1x_1 + (t_2 + t_3)\Big(\frac{t_2}{t_2 + t_3}x_2 + \frac{t_3}{t_2 + t_3}x_3\Big).$$

现在, 最右边圆括号的表达式为 S 中两个向量的凸组合, 因而也在 D 里面. 如果将它记成 d, 那么

$$t_1x_1 + t_2x_2 + t_3x_3 = t_1x_1 + (t_2 + t_3)d.$$

但是右边式子为 D 中两元素的凸组合, 故也在 D 中. 因此, 我们看到 S 中任意三个向量的所有凸组合也在 D 中. 沿着这条思路递推下去就能完成证明, 我们将此留作练习.

B.3　线性超平面和仿射超平面

我们接下来讨论 $\mathbb{R}^n$ 中的超平面. $\mathbb{R}^n$ 中的一个线性超平面是 $\mathbb{R}^n$ 中的一个 $n-1$ 维子空间. 同样地, 它是下面形式的线性方程的解集

$$a_1x_1 + \cdots + a_nx_n = 0,$$

或者

$$\langle \alpha, x\rangle = 0,$$

其中 $\alpha = (a_1, \cdots, a_n), x = (x_1, \cdots, x_n)$. 从几何上看, 就是由 $\mathbb{R}^n$ 中所有与向量 α 垂直 (正交) 的向量所形成的集合.

一个仿射超平面就是经过向量 $\beta = (b_1, \cdots, b_n)$ 变换后的线性超平面. 因此, 它是下面形式的方程的解集

$$a_1(x_1 - b_1) + \cdots + a_n(x_n - b_n) = 0,$$

或者

$$a_1x_1 + \cdots + a_nx_n = a_1b_1 + \cdots + a_nb_n,$$

或者

$$\langle \alpha, x \rangle = \langle \alpha, \beta \rangle.$$

我们用 $\mathcal{H}(\alpha, b)$ 来记超平面

$$\mathcal{H}(\alpha, b) = \{x \in \mathbb{R}^n | \langle \alpha, x \rangle = b\},$$

这里 α 是 $\mathbb{R}^n$ 中的一个向量, b 是一个实数. 注意到超平面

$$\mathcal{H}(\alpha, \| \alpha \|^2) = \{x \in \mathbb{R}^n | \langle \alpha, x \rangle = \| \alpha \|^2\}$$

包含点 α, 这是 $\mathcal{H}(\alpha, b)$ 中离原点最近的点, 因为

$$\langle \alpha, x \rangle = \| \alpha \|^2 \Rightarrow \| x \| = \| \alpha \| \Rightarrow \| x \| \geqslant \| \alpha \| .$$

一个超平面将 $\mathbb{R}^n$ 分成两个闭的半空间

$$\mathcal{H}_+(\alpha, b) = \{x \in \mathbb{R}^n | \langle \alpha, x \rangle \geqslant b\},$$
$$\mathcal{H}_-(\alpha, b) = \{x \in \mathbb{R}^n | \langle \alpha, x \rangle \leqslant b\}$$

和两个开的半空间

$$\mathcal{H}_+^o(\alpha, b) = \{x \in \mathbb{R}^n | \langle \alpha, x \rangle > b\},$$
$$\mathcal{H}_-^o(\alpha, b) = \{x \in \mathbb{R}^n | \langle \alpha, x \rangle < b\}.$$

不难证明有

$$\mathcal{H}_+(\alpha, b) \cap \mathcal{H}_-(\alpha, b) = \mathcal{H}(\alpha, b)$$

及 $\mathcal{H}_+^o(\alpha, b), \mathcal{H}_-^o(\alpha, b)$ 和 $\mathcal{H}(\alpha, b)$ 两两不相交, 且 $\mathcal{H}_+^o(\alpha, b) \cup \mathcal{H}_-^o(\alpha, b) \cup \mathcal{H}(\alpha, b) = \mathbb{R}^n$.

定义 3　称 $\mathbb{R}^n$ 中的子集 X 和 Y 被超平面 $\mathcal{H}(\alpha, b)$ 完全分离, 如果 X 位于由 $\mathcal{H}(\alpha, b)$ 确定的一个开的半空间内且 Y 位于另一半空间内. 因此, 下面陈述中必有一个成立:

1) $\langle \alpha, x \rangle < b < \langle \alpha, y \rangle$, 对所有的 $x \in X, y \in Y$.

2) $\langle \alpha, y \rangle < b < \langle \alpha, x \rangle$, 对所有的 $x \in X, y \in Y$.

B.4 分离定理

既然已经有了充分的预备知识, 我们就能开始认真考虑几个定理了. 第一个就是著名的超平面分离定理, 这是其他许多分离定理的基础.

定理 3 设 C 是 $\mathbb{R}^n$ 中的一个不包含原点的闭凸子集, 即 $0 \notin C$. 那么存在一个非零向量 $\alpha \in \mathbb{R}^n$, 使得对所有的 $x \in C$, 有

$$\langle \alpha, x \rangle \geqslant \| \alpha \|^2 .$$

因此, 超平面 $\mathcal{H}\left(\alpha, \frac{1}{2} \| \alpha \|^2\right)$ 完全分离了 0 和 C.

证明 我们首先想证明 C 包含这样的一个点, 它是 C 中所有的点中离原点最近的点. 函数

$$d(x) = \| x \|$$

是连续的, 它度量了 x 到原点的距离. 虽然 C 不一定是紧的, 但是如果能选一个实数 s, 使得球心为原点半径为 s 的闭球 $B_s(0) = \{z \in \mathbb{E}^n | \| z \| \leqslant s\}$ 与 C 相交, 那么交集

$$C' = C \cap B_s(0)$$

是闭的且有界, 故它是紧集. 因此距离函数在 C' 上可取到最小值, 不妨说在点 $\alpha \in C' \subseteq C$ 取到最小值. 我们想证明对所有的 $x \in C$, 有

$$\langle \alpha, x \rangle \geqslant \| \alpha \|^2 .$$

假设存在 $x \in C$ 满足

$$\langle \alpha, x \rangle < \| \alpha \|^2 .$$

因为 C 是凸的, 所有 α 与 x 之间的线段包含在 C 中, 即

$$\{(1-t)\alpha + tx | 0 \leqslant t \leqslant 1\} \subseteq C.$$

我们来看看该线段上一个典型的点到原点的距离. 如果发现该线段中有一点比 α 更接近原点, 那么就得到矛盾, 因为 α 是 C 中所有的点中离原点最近的点. 这就说明对所有的 $x \in C$ 有 $\langle \alpha, x \rangle \geqslant \| \alpha \|^2$, 与假设矛盾, 得证.

因此我们计算

$$\begin{aligned}\| (1-t)\alpha + tx \|^2 &= \langle (1-t)\alpha + tx, (1-t)\alpha + tx \rangle \\ &= (1-t)^2 \| \alpha \|^2 + 2t(1-t)\langle \alpha, x \rangle + t^2 \| x \|^2 \\ &= (\| \alpha \|^2 + \| x \|^2 - \langle \alpha, x \rangle)t^2 + (\langle \alpha, x \rangle - 2 \| \alpha \|^2)t + \| \alpha \|^2 .\end{aligned}$$

这是 t 的二次函数, 且向上凹, 其在下面的点上取最小值

$$t = \frac{2 \| \alpha \|^2 - \langle \alpha, x \rangle}{2(\| \alpha \|^2 + \| x \|^2 - \langle \alpha, x \rangle)}.$$

因为我们假设 $\langle\alpha,x\rangle<\|\alpha\|^2$, 又知道 $t>0$, 所以 $\|(1-t)\alpha+tx\|^2$ 的最小值严格小于 $\|\alpha\|$, 这就是我们想要证明的.

下面的结果更接近我们的目标.

定理 4 设 C 是 $\mathbb{R}^n$ 中的一个闭凸子集, 且 S 是 $\mathbb{R}^n$ 中的一个子空间, 使得 $C\cap S=\varnothing$. 那么存在一个非零向量 $\alpha\in\mathbb{R}^n$, 使得

1) $\langle\alpha,\sigma\rangle=0$, 对所有的 $\sigma\in S$(即$\alpha\in S^{\perp}$).

2) $\langle\alpha,\gamma\rangle\geqslant\|\alpha\|^2$, 对所有的 $\gamma\in C$.

因此, 超平面 $\mathcal{H}\left(\alpha,\dfrac{1}{2}\|\alpha\|^2\right)$ 完全分离了 S 和 C.

证明 考虑集合

$$A=S+C,$$

该集合是闭的, 因为 S 是闭的且 C 是紧的. 它也是凸的, 因为 S 和 C 都是凸的. 而且 $0\neq A$, 如果 $0=\sigma+\gamma$, 那么 $\gamma=-\sigma$ 就在 $C\cap S$ 里面, 而 $C\cap S$ 是空集.

故根据定理 3 可推知, 存在一个非零向量 $\alpha\in\mathbb{R}^n$, 使得对所有的 $x\in A=S+C$, 有

$$\langle\alpha,x\rangle\geqslant\|\alpha\|^2 .$$

令 $x=\sigma+\gamma$ 为 $S+C$ 上的任意一个元素, 那么

$$\langle\alpha,\sigma\rangle+\langle\alpha,\gamma\rangle=\langle\alpha,\sigma+\gamma\rangle\geqslant\|\alpha\|^2 .$$

现在, 如果对任何 $\sigma\in S$ 有 $\langle\alpha,\sigma\rangle$ 非负, 那么可用 σ 的点乘来代替 σ, 使得左边为负, 这是可以的. 因此, 对所有的 $\sigma\in S$, 一定有 $\langle\alpha,\sigma\rangle=0$, 这就是上面的 1). 又由于 $\langle\alpha,\sigma\rangle=0$, 因此也得到对所有的 $\gamma\in C$, 有

$$\langle\alpha,\gamma\rangle\geqslant\|\alpha\|^2 .$$

得证.

现在来看看我们的主要目标.

定理 5 设 S 是 $\mathbb{R}^n$ 的子空间且满足 $S\cap\mathbb{R}^n_+=\{0\}$, 其中

$$\mathbb{R}^n_+=\{(x_1,\cdots,x_n)|x_i\geqslant 0\}$$

是 $\mathbb{R}^n$ 的非负部分. 那么 $S^{\perp}$ 包含一个强正的向量.

证明 我们想要把 S 从某些地方分离开来, 但是不可能将它与 $\mathbb{R}^n_+$ 分离. 那么来考虑 $\mathbb{R}^n_+$ 的标准基 $\varepsilon_1,\cdots,\varepsilon_n$ 的凸包 Δ,

$$\Delta=\left\{t_1\varepsilon_1+\cdots+t_n\varepsilon_n|0\leqslant t_i\leqslant 1,\sum t_i=1\right\}.$$

非常明显的是 $\Delta \subseteq \mathbb{R}_+^n \backslash \{0\}$, 因此 $\Delta \cap S = \varnothing$. 同样, Δ 是闭凸集且有界, 故 Δ 是紧的. 因而, 由定理 4 知, 存在非零向量 $\alpha = (a_1, \cdots, a_n)$, 使得

1) $\alpha \in S^\perp$.

2) $\langle \alpha, \delta \rangle \geqslant \| \alpha \|^2$, 对所有的 $\delta \in \Delta$.

取 $\delta = \varepsilon_i$, 得到

$$a_i = \langle \alpha, \varepsilon_i \rangle \geqslant \| \alpha \|^2 > 0.$$

因此 α 是强正的, 得证.

下面的定理与定理 5 类似, 只是存在严格正的向量而已.

定理 6 *设 S 是 $\mathbb{R}^n$ 的子空间且满足 $S \cap \mathbb{R}_{++}^n = \{0\}$, 其中*

$$\mathbb{R}_{++}^n = \{(x_1, \cdots, x_n) | x_i > 0\}$$

是 $\mathbb{R}^n$ 的正的部分. 那么 $S^\perp$ 包含一个严格正的向量.

证明 首先注意到 $S^\perp$ 包含一个严格正的向量 α 当且仅当它包含一个分量之和为 1 的严格正的向量 (只需将 α 除以分量之和).

令 $\mathcal{B} = \{B_1, \cdots, B_k\}$ 是 S 的一个基, 考虑下面的矩阵

$$M = (B_1 | B_2 | \cdots | B_k),$$

每列为 $\mathcal{B}$ 中的基向量. 记 M 的行为 $R_1, \cdots, R_n$. 注意 $R_i \in \mathbb{R}^k$, 其中 $k = \dim(S)$.

现在, $\alpha = (a_1, \cdots, a_m) \in S^\perp$ 当且仅当 $\langle \alpha, B_i \rangle = 0, \forall i$. 这等价于矩阵方程

$$\alpha M = 0,$$

或向量方程

$$a_1 R_1 + \cdots + a_m R_m = 0.$$

因此, $S^\perp$ 包含一个分量之和为 1 的严格正的向量 $\alpha = (a_1, \cdots, a_m)$ (或等价地, 任何严格正的向量) 当且仅当

$$a_1 R_1 + \cdots + a_m R_m = 0,$$

其中系数 $a_i \geqslant 0$ 满足 $\sum a_i = 1$. 换句话说, $S^\perp$ 包含一个严格正的向量当且仅当由 $\mathbb{R}^k$ 中的 $R_1, \cdots, R_m$ 形成的凸包 C 包含 0.

现在, 我们假设 $S^\perp$ 不包含一个严格正的向量, 然后证明 S 一定包含一个强正的向量. 这样就证明了该定理. 因此, 我们假设 $0 \notin C$. 由于 C 是闭凸的, 从而由定理 3 可知, 存在一个非零向量 $\beta = (b_1, \cdots, b_k) \in \mathbb{R}^k$, 满足

$$\langle \beta, x \rangle \geqslant \| \beta \|^2 > 0, \quad \forall x \in C.$$

现在, 考虑向量

$$v = \beta_1 B_1 + \cdots + \beta_k B_k \in S \ ,$$

v 的第 i 个分量为

$$\langle \beta, R_i \rangle \geqslant \| \beta \|^2 > 0.$$

因此 v 是强正的. 这就完成了定理的证明.